

**Univerzita Karlova v Praze**

**Filozofická fakulta**

Ústav českého jazyka a teorie komunikace

**Diplomová práce**

Bc. Alena Poncarová

**Pozice antecedentu zájmena třetí osoby:  
tendence v současné češtině**

Position of Antecedent of the 3rd Person Pronoun:  
Trends in Contemporary Czech

Praha 2013

Konzultantka práce: Mgr. Pavlína Jínová  
Vedoucí práce: Mgr. Eva Lehečková, Ph.D.

## **Poděkování**

Na tomto místě bych ráda poděkovala všem, kteří mi při psaní diplomové práce byli oporou. V první řadě Mgr. Pavlíně Jínové, která mě podporovala a vždy byla ochotná cokoli konzultovat a s čímkoli poradit, bez ní by tato práce pravděpodobně nikdy nebyla hotová. Také jí vděčím za zprostředkování přístupu k RNDr. Jiřímu Mírovskému, Ph.D., který provedl všechna měření koreferenčních vztahů v PDT. Dále velice děkuji Tomáši Janečkovi, který ochotně testoval všechno, co jsem vymyslela, a odpovídal na všechny mé otázky.

Za konzultace designu výzkumu děkuji zejména Mgr. Evě Lehečkové, Ph.D. Za to, že vůbec bylo možné provést výzkum tak, jak jsme ho vymyslely, vděčím Jaroslavu Polakovičovi, který vytvořil aplikaci pro testování. Za pomoc se statistickým vyhodnocením moc děkuji Michalovi Nguyen Quang Dungovi, který mi věnoval velké množství volného času a energie, a Mgr. Janu Chromému, Ph.D., který se mnou znovu a znovu procházel materiály a navrhoval nová a lepší řešení.

Velkým přínosem byli i Jana Česká, která naslouchala vždy, když jsem potřebovala, a pomáhala mi se vším, co bylo potřeba, Jiří Pergler, který se mnou zasvěceně diskutoval, Helena Filipová, která objektivně hodnotila a konstruktivně kritizovala, a Anna Morávková a Karolína Vyskočilová, které byly nekonečným zdrojem materiálů. Za pomoc při distribuci dotazníků vděčím Janě České, Tomáši Janečkovi, všem Poncarovým, Plachým a Kukrálovým a dalším a dalším.

Děkuji Vám všem!

V Praze dne 20. 7. 2013

Alena Poncarová

### **Prohlášení**

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů a že tato práce nebyla použita k získání žádné jiné atestace ani titulu.

V Praze dne 20. 7. 2013

Alena Poncarová

UNIVERZITA KARLOVA V PRAZE  
Filozofická fakulta

Ústav českého jazyka a teorie komunikace

Akademický rok: 2011/2012

ZADÁNÍ DIPLOMOVÉ PRÁCE

Jméno a příjmení: **Alena Poncarová**

studijní program: **Filologie**

studijní obor: **Český jazyk - specializační studium**

Děkan fakulty Vám podle zákona č. 111/1998 Sb. určuje tuto diplomovou práci:

Název práce: **Pozice antecedentu zájmena třetí osoby: tendence v současné češtině**

Anglický název práce: **Position of Antecedent of the 3rd Person Pronoun: Trends in Contemporary Czech**

Jazyk práce: **čeština**

Zásady pro vypracování:

Hlavním cílem práce je popsat tendence umisťování antecedentu zájmena třetí osoby v textu. Tvary zájmena 3. osoby se přitom budou rozumět nejen tvary přítomné v povrchové struktuře, ale rovněž jeho formy nulové. Tendence v umisťování antecedentu zájmena třetí osoby budou v práci zkoumány dvojím způsobem: Nejprve bude analyzován vzorek současných českých textů (předpokládá se využití publicistických textů) s cílem zjistit a) jak často je antecedentu zájmena třetí osoby umístěn v tématu a jak často v rématu některé z předcházejících výpovědí, b) zda se vyskytují preference v navazování na některé větné členy ve srovnání s jinými (např. navazuje se na podmět předchozí věty častěji než na její jiné větné členy?), c) jak daleko (v počtu výpovědí) v předchozím kontextu se antecedent nachází.

Druhým způsobem zkoumání tendencí v umisťování antecedentu zájmena třetí osoby bude experimentální situace: rodilí mluvčí češtiny budou vyjadřovat svoje preferované interpretace v kontextech, kdy antecedentu zájmena třetí osoby nebude zcela jasný z kontextu (*Mirek přišel za Lukášem na návštěvu. Ukazoval mu schopnosti své nové flashky.*). Z tohoto experimentu vyplyne, které faktory jsou pro určování antecedentu pro recipienty podstatnější než jiné. Zároveň se očekává korespondence s nálezy v textech.

Teoretickými základy práce je koncept tematických posloupností (Daneš, 1968) a teorie centeringu (Grosz, Weinstein, Joshi, 1995). Teorie centeringu byla zvolena proto, že přináší predikce umisťování antecedentů zájmen, které budou využity pro analýzu českých dat.

Seznam odborné literatury:

- HAIČOVÁ, Eva – PARTEE, Barbara – SGALL, Petr: *Topic, focus articulation, tripartite structures and semantic content*. Dordrecht: Kluwer Academic Press, 1998.
- DANEŠ, František: Typy tematických posloupností v textu (na materiále českého textu odborného). *SaS* 29/1968, s. 125 - 141. DANEŠ, František (ed.): *Papers on functional sentence perspective*. Praha: Academia, 1974. Vydání první. DANEŠ, František - HLAVSA, Zdeněk - KORENSKÝ, Jan: *Práce o sémantické struktuře věty*. Praha: ÚJČ AV ČSAV, 1973. Vydání první.
- GROSZ, Barbara - WEINSTEIN, Scott – JOSHI, Aravind: Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics* 2/1995, s. 203 – 225.
- HOFFMANNOVÁ, Jana: *Sémantické a pragmatické aspekty koherence textu*. Praha: ÚJČ AV ČR, 1983.
- MIKULOVÁ, Marie et al.: *Anotace na tektogramatické rovině Pražského závislostního korpusu. Anotátorská příručka*. Praha: UFAL MFF, 2005.
- NĚDOLUŽKO, Anna: *Rozšířená textová koreference a asociační anafora*. Praha: ÚFAL MFF, 2011.
- NOVÁK, Pavel: O prostředcích aktuálního členění. *Acta Universitatis Carolinae - Philologica* 1/1959, s. 9 - 15.
- PANEVOVÁ, Jarmila: K významové stavbě větného centra. *SaS* 39/1978, s. 140 - 148.
- SGALL, Petr: Základ a jádro, nebo presupozice a ohnisko? *SaS* 33/1972, s. 160 - 163.
- SGALL, Petr: K obecným otázkám sémantiky věty. *SaS* 37/1976, s. 184 - 194.
- STEELE, Ross - THREADGOLD, Terry: *Language Topics - Essays In Honour Of Michael Halliday*. Volume 1, 2. Amsterdam, Philadelphia: John Benjamins, 1987.
- TRÁVNÍČEK, František: O tak zvaném aktuálním členění větném. *SaS* 22/1961, s. 163 - 171.
- WALKER, Marilyn - JOSHI, Aravind - FRIEDMAN PRINCE, Ellen (eds.): *Centering Theory in Discourse*. Oxford: Clarendon Press, 1998. First Edition.

Vedoucí diplomové práce: **Mgr. Lehečková Eva, Ph.D.**



Konzultanti: **Mgr. Jínová Pavlína**

Datum zadání diplomové práce: 24.5.2012

Termín odevzdání diplomové práce: dle harmonogramu příslušného akademického roku



Vedoucí katedry



Děkan

V Praze dne 6.6.2012

## **Klíčová slova**

aktuální členění

anafor

antecedent

koreference

podmět

postcedent

předmět

réma

téma

zájmeno

## **Key words**

anafor

antecedent

coreference

information structure

focus

object

postcedent

pronoun

subject

topic

**Abstrakt**

*Cílem této diplomové práce je zjistit, jakým způsobem probíhá navazování koreferenčních vztahů v českém textu z hlediska aktuálního členění a větněčlenské platnosti. Ve zpracování této problematiky jsem postupovala následujícím způsobem.*

*Prozkoumala jsem dostupnou literaturu týkající se aktuálního členění a větněčlenské platnosti s ohledem na utváření koreferenčních vztahů a její poznatky jsem poté konfrontovala s vlastním výzkumem. Vzhledem k nejednotnosti přístupů bylo nejprve nutné představit termíny, s nimiž tato diplomová práce operuje (kapitola 2) a poté bylo možné představit dosavadní přístupy k problematice včetně metodologického základu mé práce, který představuje zahraniční teorie centeringu, protože, na rozdíl od českých přístupů, poskytuje predikce o četnostech různých způsobů navazování (kapitola 3).*

*Poznatky získané z literatury byly poté aplikovány na data dvojího typu, korpusový materiál Pražského závislostního korpusu a materiál získaný prostřednictvím dotazníkového výzkumu (kapitola 4). Prostřednictvím konfrontace predikcí s daty byly stanoveny nejčastější způsoby navazování koreferenčních řetězců v českém textu, získali jsme tedy informace o principech výstavby českého textu z hlediska uspořádání aktuálně- a větněčlenského (kapitola 5).*

**Abstract**

*The aim of this thesis is to find out how co-referential relationships in Czech texts are developed in terms of the information and constituent structures. The data was processed using the following methods.*

*Firstly I examined the available literature on information and constituent structures with special attention being paid to the formation of coreferential relations. This information was then compared to my own research. With regards to disparities between the approaches, it was necessary to introduce some basic terms that are used within my thesis (Chapter 2). Thereafter, I was able to introduce current approaches to the thesis, including the methodological basis of my work, which is represented by the foreign Centering Theory, because unlike other Czech approaches this theory includes predictions on the frequency of different ways of co-referential relationship development (Chapter 3).*

*Secondly, the knowledge gained from literature was applied to two different types of data: corpus material of Prague Dependency Treebank and material obtained through a questionnaire survey (Chapter 4). The most appropriate methods of developing co-referential chains in the Czech text were determined by challenging the prediction model with the obtained data set. Using this method we gained information about structural principles of Czech text – from the information and constituent point of view. (Chapter 5).*

## Obsah

<b>1</b>	<b>ÚVOD.....</b>	<b>12</b>
<b>2</b>	<b>KOREFERENCE, AKTUÁLNÍ ČLENĚNÍ A VĚTNÉ ČLENY .....</b>	<b>14</b>
<b>2.1</b>	<b>KOREFERENCE.....</b>	<b>15</b>
2.1.1	<i>Antecedent .....</i>	<i>19</i>
2.1.2	<i>Anafor .....</i>	<i>19</i>
2.1.3	<i>Zájmeno třetí osoby .....</i>	<i>20</i>
<b>2.2</b>	<b>AKTUÁLNÍ ČLENĚNÍ .....</b>	<b>21</b>
2.2.1	<i>Kontextová zapojenost .....</i>	<i>23</i>
2.2.2	<i>Výpovědní dynamičnost.....</i>	<i>26</i>
2.2.3	<i>Topik .....</i>	<i>29</i>
2.2.4	<i>Fokus.....</i>	<i>30</i>
<b>2.3</b>	<b>VĚTNÝ ČLEN .....</b>	<b>30</b>
2.3.1	<i>Podmět .....</i>	<i>31</i>
2.3.2	<i>Předmět .....</i>	<i>31</i>
<b>3</b>	<b>DOSAVADNÍ PŘÍSTUPY .....</b>	<b>33</b>
<b>3.1</b>	<b>VILÉM MATHESIUS .....</b>	<b>33</b>
<b>3.2</b>	<b>FRANTIŠEK DANEŠ.....</b>	<b>33</b>
3.2.1	<i>Návazná tematizace rématu .....</i>	<i>37</i>
3.2.2	<i>Průběžné téma .....</i>	<i>38</i>
3.2.3	<i>Rozvíjení rozštěpeného rématu.....</i>	<i>39</i>
3.2.4	<i>Tematické odvozování.....</i>	<i>40</i>
3.2.5	<i>Tematický skok.....</i>	<i>41</i>
<b>3.3</b>	<b>LUDMILA ZIMOVÁ .....</b>	<b>43</b>
3.3.1	<i>Podmět .....</i>	<i>43</i>
3.3.2	<i>Předmět .....</i>	<i>44</i>
3.3.3	<i>Konkurence vyjadřovacích prostředků .....</i>	<i>47</i>
3.3.4	<i>Topik – fokus .....</i>	<i>48</i>
<b>3.4</b>	<b>CENTERING.....</b>	<b>50</b>
3.4.1	<i>Centra pozornosti .....</i>	<i>50</i>
3.4.2	<i>Typy přechodů .....</i>	<i>53</i>



3.4.3	<i>Zájmena</i> .....	57
3.4.4	<i>Hypotézy o navazování koreferenčních řetězců</i> .....	59
<b>4</b>	<b>ANTECEDENT ZÁJMENA TŘETÍ OSOBY</b> .....	<b>61</b>
<b>4.1</b>	<b>PDT</b> .....	<b>61</b>
4.1.1	<i>Aktuální členění v PDT</i> .....	61
4.1.2	<i>Koreference v PDT</i> .....	63
4.1.3	<i>Zkoumání koreference</i> .....	66
4.1.4	<i>Výsledky</i> .....	68
4.1.5	<i>Shrnutí</i> .....	73
<b>4.2</b>	<b>DOTAZNÍKOVÝ VÝZKUM</b> .....	<b>75</b>
4.2.1	<i>Cíl výzkumu</i> .....	75
4.2.2	<i>Metoda výzkumu</i> .....	75
4.2.3	<i>Proměnné</i> .....	76
4.2.4	<i>Testovací kontexty</i> .....	79
4.2.5	<i>Zájmenná koreference</i> .....	82
4.2.6	<i>Nejednoznačné navazování</i> .....	87
4.2.7	<i>Testování přijatelnosti</i> .....	89
4.2.8	<i>Výplňkové kontexty</i> .....	95
4.2.9	<i>Testování navazování koreferenčních řetězců</i> .....	96
<b>4.3</b>	<b>VÝSLEDKY DOTAZNÍKOVÉHO VÝZKUMU</b> .....	<b>104</b>
4.3.1	<i>Aktuální členění</i> .....	106
4.3.2	<i>Větněčlenská platnost</i> .....	108
4.3.3	<i>Jednotlivé kontexty</i> .....	111
4.3.4	<i>Zájmena</i> .....	114
4.3.5	<i>Kritická místa výzkumu</i> .....	115
<b>5</b>	<b>ZÁVĚR</b> .....	<b>117</b>
<b>5.1</b>	<b>UTVÁŘENÍ KOREFERENČNÍCH ŘETĚZCŮ RECIPIENTEM TEXTU</b> .....	<b>117</b>
<b>5.2</b>	<b>CENTERINGOVÁ TEORIE A ČEŠTINA</b> .....	<b>119</b>
<b>5.3</b>	<b>REFERENČNÍ NESPECIFIKOVANOST</b> .....	<b>120</b>
<b>5.4</b>	<b>SÍLA PROMĚNNÝCH</b> .....	<b>121</b>
<b>5.5</b>	<b>SLABINY DOSAVADNÍCH PŘÍSTUPŮ</b> .....	<b>122</b>
<b>6</b>	<b>SEZNAM LITERATURY</b> .....	<b>123</b>

<b>7</b>	<b>SEZNAM SCHÉMAT .....</b>	<b>127</b>
<b>8</b>	<b>SEZNAM TABULEK .....</b>	<b>132</b>
<b>9</b>	<b>SEZNAM PŘÍKLADŮ .....</b>	<b>144</b>
<b>10</b>	<b>PŘÍLOHY .....</b>	<b>146</b>
<b>10.1</b>	<b>ALGORITMUS A PŘÍKLAD (BFP 1987: 159) .....</b>	<b>147</b>
<b>10.2</b>	<b>KONTEXTY TESTOVÁNÍ PŘIJATELNOSTI.....</b>	<b>148</b>
<b>10.3</b>	<b>VÝSLEDKY TESTOVÁNÍ PŘIJATELNOSTI .....</b>	<b>150</b>
<b>10.4</b>	<b>SLEDOVANÉ POLOŽKY .....</b>	<b>158</b>
<b>10.5</b>	<b>T–TEST .....</b>	<b>161</b>
<b>10.6</b>	<b>DOTAZNÍKOVÉ POLOŽKY .....</b>	<b>162</b>
<b>10.7</b>	<b>OTÁZKY A MOŽNOSTI ODPOVĚDI .....</b>	<b>167</b>
<b>10.8</b>	<b>T–F PODLE NASTAVENÍ KONTEXTU (CHI-SQUARE, P-VALUES).....</b>	<b>168</b>
<b>10.9</b>	<b>PO–PT PODLE NASTAVENÍ KONTEXTU (CHI-SQUARE, P-VALUES) .....</b>	<b>168</b>

**Seznam použitých zkratk**

AČV	aktuální členění větné
Ag	agens
BFP	Brennan – Friedman – Pollard (1987)
Cb	centrum pozornosti typu Cb ( <i>backward-looking center</i> )
Cf	centrum pozornosti typu Cf ( <i>forward-looking center</i> )
Cp	centrum pozornosti typu Cp ( <i>preferred center</i> )
F	fokus
<i>i</i>	označení koreferenčního řetězce
<i>i1/i2</i>	nejednoznačný koreferenční řetězec
N1	první jmenná fráze v dotazníkovém kontextu
N2	druhá jmenná fráze v dotazníkovém kontextu
P	zájmeno
PDT	Pražský závislostní korpus ( <i>Prague Dependency Treebank</i> )
Po	podmět
Pt	předmět
PRO	generované (povrchově nevyjádřené) zájmeno
Re	referující výraz ( <i>referring expression</i> )
T	topik
TFA	topic-focus articulation
tfa	atribut tfa v PDT
$U_i$	aktuální výpověď
$U_{i-1}$	bezprostředně předcházející výpověď
VČ	větný člen, větněčlenská platnost
WJP	Walker – Joshi – Prince (1998)
X	další doplnění, další větné členy (nespecifikováno)

# 1 Úvod

Základem úspěšné komunikace je porozumění smyslu textu, tedy porozumění významům jednotlivých slov i vztahům mezi nimi. Potřeba zkoumat, jakým způsobem jsou mluvčí schopni dešifrovat princip výstavby textu, vyvozovat jeho význam a interpretovat smysl, má mnoho aspektů. Jednomu z nich, konkrétně koreferenci, bude věnována tato práce.

Téma výstavby textu bývá tradičně řazeno do oblasti textové lingvistiky a syntaxe, my se mu budeme věnovat z hlediska koreference. Perspektivu textově lingvistickou představuje především aktuální členění, perspektivu syntaktickou větněčlenská platnost.

Tyto dva přístupy ke koreferenci zastupují hlavní linii zkoumání propojující práce české lingvistiky na straně jedné v čele s Mathesiem a Danešem (v případě aktuálního členění) a na straně druhé Zimovou (v případě větněčlenské platnosti) s metodologií zahraniční teorie centeringu (kde se jedná především o práce Brennanové, Groszové, Joshiho, Princeové, Sidnerové a Walkerové).

Primárním východiskem zkoumání navazování koreferenčních řetězců je americká syntaktická teorie centeringu, která pracuje s aktuálním členěním a větněčlenským postavením. Považujeme za vhodné představit teorii aktuálního členění v rámci českého kontextu (Daneš, Mathesius, Sgall, Hajičová) zejména proto, že americký přístup centeringu s ním plně nekoresponduje, tudíž je žádoucí vymezit jejich vzájemný vztah. Aspekty větněčlenské platnosti nejsou v českém prostředí od amerických odlišné, nicméně budou představeny také, aby byl výklad ucelený.

Akademická Mluvnice češtiny 3 (1987) vymezuje dva základní přístupy k výstavbě textu. Prvním z nich jsou kognitivní významové složky větného významu vyplývající z propozice a vymezení sémantických participantů, druhým jsou tzv. komunikativně-pragmatické složky, kam patří právě aktuální členění a procesy jako například aktualizace propozice a specifikace participantů z hlediska denotace.

Textová lingvistika má v české tradici pevné místo již od Mathesia (1939). Za jeden ze základních textů zabývajících se výstavbou koreferenčních řetězců považujeme Danešův příspěvek o tematických posloupnostech (1968), proto z jeho klasifikace výstavby textu budeme vycházet. Dále se zaměříme na práce Ludmily Zimové, zejména *Způsoby vyjadřování větných členů v textu* (1994), která se zabývá vyjadřováním větných členů v textu. V českém prostředí dosud nijak výrazně reflektovaným teoretickým

východiskem bude syntaktická teorie centeringu (např. WJP 1998), která se zaměřuje na zkoumání lokální koherence textu.

Jazyková data, která poslouží jako východisko zkoumání, budeme čerpat především ze dvou zdrojů. Prvním z nich je Pražský závislostní korpus, v němž jsou koreferenční vztahy anotovány (Mikulová et al. 2005), a druhým dotazníkový výzkum, který jsme provedli na přelomu let 2012 a 2013. Hlavní část diplomové práce je věnována dotazníkovému šetření navazování koreferenčních řetězců v českých textech.

## 2 Koreference, aktuální členění a větné členy

Objektem našeho zájmu je vztah koreference s aktuálním členěním a větněčlenskou platností. V následující části textu se budeme věnovat potřebným terminologickým východiskům, definujeme zájmeno třetí osoby, koreferenci a v jejím rámci antecedent a anafor, aktuální členění a v jeho rámci základ, východisko, jádro, ohnisko, téma a réma, topik a fokus.

Vzhledem k metodě, kterou aplikujeme v dotazníkovém výzkumu, je nutné představit pojmy, s nimiž operujeme. Zdá se to být přínosné z toho důvodu, že terminologie není jednotná, srovnajme například pojmy *téma*, *réma* a *topik*, *fokus* nebo *ohnisko* a *jádro*, liší se zejména tradice česká mathesiovská a danešovská od anglické. Problém nejednotnosti terminologie tematizoval například Petr Sgall (1972).

V důsledku faktu, že se zabýváme velice specifickou oblastí textové lingvistiky a důraz klademe na koreferenční vztahy v textu v rámci omezeného kontextu, nerozlišujeme některé terminologické distinkce, které bývají uplatňovány. Jde zejména o dokulilovskou (Dokulil 1958: 109) distinkci věty a výpovědi (preferujeme užívání termínu výpověď)<sup>1</sup> a distinkci termínů kontext a kotext (používám jen termín kontext pro jazykové i mimojazykové okolí). Naopak distinkce mezi denotátem a referentem<sup>2</sup> je distinkcí, kterou aplikujeme důsledně.

Nyní přistoupíme k definici první skupiny termínů, tedy termínů koreference, antecedent, anafor a zájmeno třetí osoby.

<sup>1</sup> Termínu výpověď užíváme v kapitole 2 proto, že věta je v tradičním pojetí jednotkou abstraktní, systémovou, a tudíž termín výpověď adekvátněji postihuje konkrétní jazykové vyjádření, s nímž pracujeme. Termínu věta je užíváno zejména v částech týkajících se prací Ludmily Zimové, neboť ta se zaměřuje na zkoumání větných členů, a dále v některých částech týkajících se dotazníkového šetření, neboť tam odhlížíme od konkrétní realizace výpovědní.

<sup>2</sup> Referent chápeme v souladu s Čermákovým pojetím jako vztah mezi označením, tedy jazykovým vyjádřením, a mimojazykovou skutečností, tedy předmětem, který je konkrétní a který je zástupcem dané skupiny. „Reference: bilaterální (a momentální) vztah znaku [...] v kontextu k *individuálnímu* objektu [...], tj. k **referentu**; je typicky textové povahy.“ (Čermák 2010: 105) Denotát je také označením vztahu mezi jazykovým vyjádřením a mimojazykovou skutečností, ale denotát na rozdíl od referentu není konkrétní, referuje ke třídě, k typu, ke skupině referentů: „Denotace: bilaterální vztah znaku ke *třídě* objektů (i abstraktních), **denotátu**, který má typicky slovníkovou povahu.“ (Čermák 2010: 104)

## 2.1 Koreference

Vztah koreference je definován jako „referenční shoda výrazů v textu (tj. totožnost jejich referentů, entit, ke kterým odkazují)“ (ESČ 2002: 233). Kritériem koreference je tedy mimojazyková skutečnost, uspořádání reality. Základní otázkou je, zda dvě vyjádření v textu odkazují ke stejné entitě skutečnosti, zda mají dvě vyjádření v textu shodnou referenci. Přesto je koreference považována za vztah vnitrotextový, endoforický (v opozici k exoforickému, který referuje mimo text sám), neboť jde vždy o vztah dvou textových vyjádření.

Vedle termínu koreference bývá často užíváno i termínu izotopie (např. MČ III 1987: 633); Ludmila Zimová v rámci své významné práce věnované vyjadřování větných členů v textu (1994) používá termínu izotopický řetězec. Vztah izotopie je definován jako „významová souvislost lexikálních jednotek v (po sobě následujících) segmentech textu, která se utváří na základě opakování lexémů nebo sémémů“ (ESČ 2002: 190). Je zřejmé, že termíny koreference a izotopie nejsou zcela synonymní. Koreference odkazuje ke shodě referentů, tedy pohybuje se na úrovni denotační vs. referenční, izotopie oproti tomu odkazuje k parataktickému vztahu mezi jednotlivými výrazy, přičemž tímto vztahem nemusí být nutně totožnost.

My pracujeme s koreferencí ve smyslu její definice v ESČ, z toho důvodu nepoužíváme označení izotopický řetězec, protože nekonzistence označování řetězce jako izotopického s termínem koreference by do této práce vnesla terminologickou nejednotnost. Užívanými pojmy tedy budou: koreference, koreferenční řetězec a jeho členy apod. Např. v kontextu typu *Babička slavila minulý víkend narozeniny. Darovali jsme jí pobyt v lázních.* představují výrazy *babička* a *jí* členy koreferenčního řetězce a jsou v tomto kontextu nositeli vztahu koreference. Obecně, bez důrazu na to, které členy do koreferenčního vztahu vstupují, je k nim referováno jako ke koreferenčnímu řetězci.

V tradičním pojetí vymezujeme koreferenci gramatickou a textovou (ESČ 2002: 233, Mikulová et al. 2005: 128n.). Gramatickou koreferenci charakterizuje zásada, že „se užití jazykově koreferenčních prostředků řídí gramatickými pravidly“ (ibid: 233), takováto koreference je tedy signalizována například kongruentními prostředky (jako jsou např. zájmena vztažná nebo zájmena reflexivní) a agenty infinitivních tvarů v konstrukcích s tzv. slovesy kontroly. Oproti tomu je za textovou koreferenci označováno „užití takových

jazykových prostředků, které odkazují k objektu již zmíněnému (anafora) nebo (řidčeji) k objektu, který teprve bude v diskurzu instalován (katafora)“ (ESČ 2002: 234). Textová koreference je tedy utvářena na úrovni samotných lexikálních jednotek, ne v rámci jejich struktury. Zároveň není v případě textové koreference repertoár prostředků, které mohou vytvářet textově koreferenční řetězec, omezen na některé typy vyjádření, nejčastěji bývá realizován opakováním lexikálního pojmenování, pronominalizací nebo elizí. Prostředky, které bývají uváděny jako prototypičtí zástupci prostředků signalizujících vztahy gramatické koreference, mohou sloužit i koreferenci textové, hranice mezi těmito dvěma typy koreference není v žádném případě jasná a jednoznačná. Předmětem našeho zkoumání jsou případy, kdy není koreferenční řetězec jasně specifikován, přestože máme k dispozici různé ukazatele, které koreferenci naznačují. Tato nespecifikovanost koreferenčního řetězce je způsobena tím, že v kontextu je k dispozici více potenciálních antecedentů, z nichž ani jeden není pomocí gramatických pravidel jazyka vyloučen. Více viz oddíl Testovací kontexty (4.2.4).

Členy koreferenčního řetězce, tj. ta vyjádření, která referují ke stejné entitě mimojazykové skutečnosti, jsou při analýze tradičně označovány koindexací (v případě této práce používáme koindexaci *i* doplněnou o číselné rozlišení, pokud se prolíná několik koreferenčních řetězců).

K utváření koreferenčního řetězce je možno z hlediska toho, co je jeho hlavním konstitutivním prvkem, přistupovat dvěma základními způsoby uplatňovanými (též) v teorii centeringu. První možný přístup vychází z předpokladu, že každé pojmenování v textu otevírá možnost utvoření koreferenčního řetězce. Takovýto přístup bychom mohli nazvat *forward-looking*. Oproti tomu přístup, který nazýváme *backward-looking*, pracuje s předpokladem, že některá vyjádření v textu (prototypicky zájmena a deiktika obecně) v předchozím textu „vyhledávají“ potenciální antecedenty, k nimž by mohla být prostřednictvím koreferenčního řetězce navázána. V této práci je upřednostňován přístup *backward-looking*, neboť se shoduje s perspektivou, kterou zaujímáme v dotazníkovém šetření (viz kapitola Dotazníkový výzkum – 4.2), tj. zájmeno si volí svůj antecedent.

Při zkoumání koreference musíme uvažovat o tom, v rámci jakého rozsahu textu budeme koreferenci zkoumat. Existují dvě základní možnosti, první z nich je zkoumat členy v rámci výpovědi (typicky u vztahu gramatické koreference, např. reflexivního



zájmena: *Petr si koupil nové auto.*), druhá se zaměřuje na větší části textu, na vztahy překračující hranici jedné výpovědi, tedy na koreferenci textovou (*Petr má moc rád svoji maminku. Petr/on/0 mamince/jí koupil nové auto.*). Perspektivu větnou aplikuje analýza větných členů, perspektivu textovou analýza větného členění. Existují různé termíny, kterými se tato dvě členění označují. Mathesius (1947) staví větné členění, které označuje jako formální, do protikladu k členění aktuálnímu, Trávníček (1961) k Mathesiovým termínům vztahuje termíny větného a členského slovosledu oproti slovosledu významovému (tyto termíny nacházíme i v Trávníčkově Mluvnici spisovné češtiny). Přehledně a zjednodušeně je Mathesiovo pojetí znázorněno v tabulce A, v níž jsou rozlišeny vztahy gramatické struktury k věcné souvislosti a gramatického subjektu, resp. predikátu, k východišti, resp. jádru výpovědi.

<b>formální členění</b>	<b>aktuální členění</b>
gramatická struktura	věcná souvislost
gramatický subjekt gramatický predikát	východiště výpovědi jádro výpovědi

**Tabulka A: formální a aktuální členění**

S rozsahem koreferovaných členů souvisí i příspěvek Rolanda Harwega z roku 1994, který se zabývá otázkou, zda v případě pronominalizovaných anaforů nominálních nebo pronominálních vyjádření dochází k inkluzi predikátů antecedentních vět, či nikoli. Harweg analyzuje několik typů pronominalizovaných vyjádření z hlediska zahrnutí predikátů antecedentních, které se navzájem odlišují, nicméně uvádí, že „fakt, že [pronominalizovaná vyjádření] nezahrnují predikáty svých antecedentů, implikuje, že, v přesném slova smyslu, nemohou být interpretována jako *anaforická*, nebo minimálně ne jako koreferenční anaforická vyjádření těchto antecedentů.“<sup>3</sup> Pokud tedy anafor nezahrnuje predikát svého antecedentu, nepovažuje ho Harweg za koreferenční člen daného řetězce. Harweg uvádí následující příklad, který podle něj z uvedených důvodů nelze považovat za koreferenční vztah, neboť vztah mezi predikátem antecedentního vyjádření (*dostal*) a

<sup>3</sup> „The fact that they detach the predicates of their antecedents from these very antecedents implies, however, that, strictly speaking, they cannot be interpreted as *anaphorics* or at least, not as coreferential anaphorics of these antecedents.“ (Harweg 1994: 143)

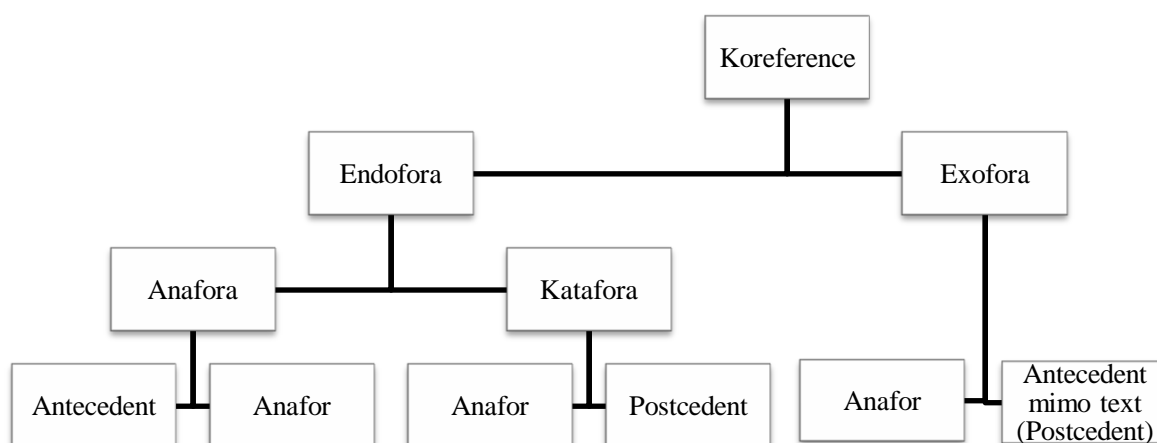
antecedentem (*dopis*) není přenášen do výpovědi B; vyjádření *dopis* ve výpovědi B odkazuje k antecedentu *dopis* v A, nikoli k situaci v A *dostat dopis*<sup>4</sup>:

*A: Včera jsem dostal zvláštní dopis. Obsahoval nabídku záhadného podnikání, kterým by si měl člověk být schopen vydělat velké množství peněz v krátkém čase.*

*B: Také jsem takový dopis dostal.*

Harwegův článek nás přivádí k uvažování o tom, zda je navazování koreferenčních vazeb v textu ovlivněno vztahem členů koreferenčního řetězce k predikátu (inkluzivitou a exkluzivitou). Tento aspekt (členění antecedentů na inkluzivní a exkluzivní) ve smyslu (ne)zahrnutí antecedentního predikátu do koreferenčního řetězce v našem výzkumu neprovádíme, protože se zaměřujeme na vztahy mezi jmennými frázemi bez jejich dalšího vnitřního dělení a členění, nicméně pro další výzkum je to jistě inspirativní přístup.

Nyní přistoupíme k definici termínů koreference, s nimiž v této budeme pracovat. Jde o termíny antecedent, anafor a zájmeno třetí osoby jakožto jeden z typů anaforu. Jejich vzájemné vztahy zachycuje schéma A. Typ vztahu, kterým se zabýváme, by v rámci uvedeného schématu A patřil do části anaforické, jde tedy o endoforické odkazování, v němž antecedent předchází svůj anafor.



**Schéma A: koreference**

<sup>4</sup> Harweg (1994:143) uvádí příklad v angličtině, překlad AP.

Vzhledem k tomu, že ostatní typy navazování z uvedeného schématu nejsou v této práci předmětem zkoumání, odkazují dále užívané termíny jako koreference, koreferenční výrazy, koreferenční vztahy v textu apod. vždy k tomuto specifickému typu vztahu endoforickému–anaforickému, pokud není specifikováno jinak.

### *2.1.1 Antecedent*

Nejprve představíme definici pojmu antecedent. Klíčovým je tento pojem z toho důvodu, že zkoumání koreferenčních řetězců vyžaduje nejprve stanovení prvků, které koreferenční řetězec tvoří (proces určování koreferujících členů, resp. antecedentu, je předmětem dotazníkového výzkumu).

Za antecedent považujeme „element poskytující referenci endofoře: je-li antecedent v textu před endoforou, jde o anaforu [...], je-li antecedent v textu po endofoře, jde o kataforu“ (ESČ 2002: 41). V tomto smyslu užíváme termínu antecedent i v této práci, omezuje se ale pouze na případy, kdy antecedent předchází endofoře, tedy v terminologii ESČ na anaforu. Termínu anafora bývá ale užíváno „i jako synonyma s nadřazeným termínem endofora“ (ibid), což může být poněkud matoucí. Vzhledem k tomu, že pracuji pouze s jedním typem koreferenční situace, není pro tuto práci distinkce endofora–anafora důležitá. S anaforou je v jiném smyslu pracováno i v PDT (4.1).

### *2.1.2 Anafor*

Termínu anafor užíváme zejména jako protikladu termínu antecedent. Jestliže antecedent označuje element, který poskytuje referenci, tedy výraz, ke kterému se odkazuje, za anafor označujeme ten element, který referenci navazuje, vyhledává, tj. jde o výraz, který odkazuje.

Ludmila Zimová pro označení anaforu používá termín postcedent, který se nám ale nezdá být vhodný, neboť ho chápeme jako protiklad antecedentu anafory ke katafoře – anaforické prostředky odkazují k antecedentům, kataforické k postcedentům. Pro takovýto způsob interpretace termínu postcedent svědčí i fakt, že termíny ante-cedent a post-cedent jsou stejně formálně utvořeny, a pokud za antecedent považujeme označení referenta odkazovacího prostředku (např. zájmena) v předchozím kontextu, logicky bychom za postcedent měli považovat označení referenta odkazovacího prostředku v následujícím

kontextu. Zimová v zásadě chápe termín postcedent na základě stejného principu, jen její perspektiva je jiná a v jejím důsledku i aplikace. Termínu antecedent užívá k označení části koreferenčního řetězce, na kterou se odkazuje, zatímco postcedent užívá jako označení té části koreferenčního řetězce, která odkazuje. Srovnání postcedentu jako části textu, na kterou se odkazuje a která je až za odkazujícím koreferenčním prostředkem, s chápáním postcedentu jako části textu, jejímž prostřednictvím se odkazuje, zřetelně ukazuje odlišnost perspektiv. V našem případě termín postcedent chápeme jako označení pro prvek podobný antecedentu, ale omezený na kataforické odkazování, pro prvek koreferenčního řetězce, který v rámci anaforického odkazování tvoří část odkazující (nikoli tu, na kterou je odkazováno, tedy antecedent), užíváme označení anafor. Volba tohoto pojetí je v zásadě arbitrární, ale bylo nutné tuto terminologickou distinkci vyjasnit, aby bylo jednoznačné, jakým způsobem s termíny pracujeme.

### *2.1.3 Zájmeno třetí osoby*

Dalším termínem, který v rámci koreference definujeme, je zájmeno třetí osoby. Potřeba definovat zájmeno třetí osoby vychází ze zaměření a omezení zkoumání koreferenčních řetězců na takový typ navazování vztahů v textu, v němž jako anafory v koreferenčních řetězcích fungují zájmena. Zájmeno třetí osoby je v našem pojetí považováno za prototypického reprezentanta členu koreferenčního řetězce v pozici anaforu. S pojmem zájmeno třetí osoby pracujeme jako se zastřešujícím pojmem pro všechny tvary zájmena *on* v obou číslech (*on, oni*) a ve všech rodech (*on, ona, ono*). Důležitým faktorem výzkumu je omezení týkající se výběru zájmen, který je blíže popsán v oddíle Zájmenná koreference (4.2.5). Na tomto místě postačí zmínit, že se neomezujeme jen na zájmena osobní, ale pracujeme i se zájmeny přivlastňovacími vztahujícími se ke třetí osobě.

Po koreferenci jakožto obecném tématu práce se zaměříme na dvě perspektivy, které jsou při nahlížení na koreferenci uplatňovány. První z nich je aktuální členění, druhou perspektiva větněčlenská.

## 2.2 Aktuální členění

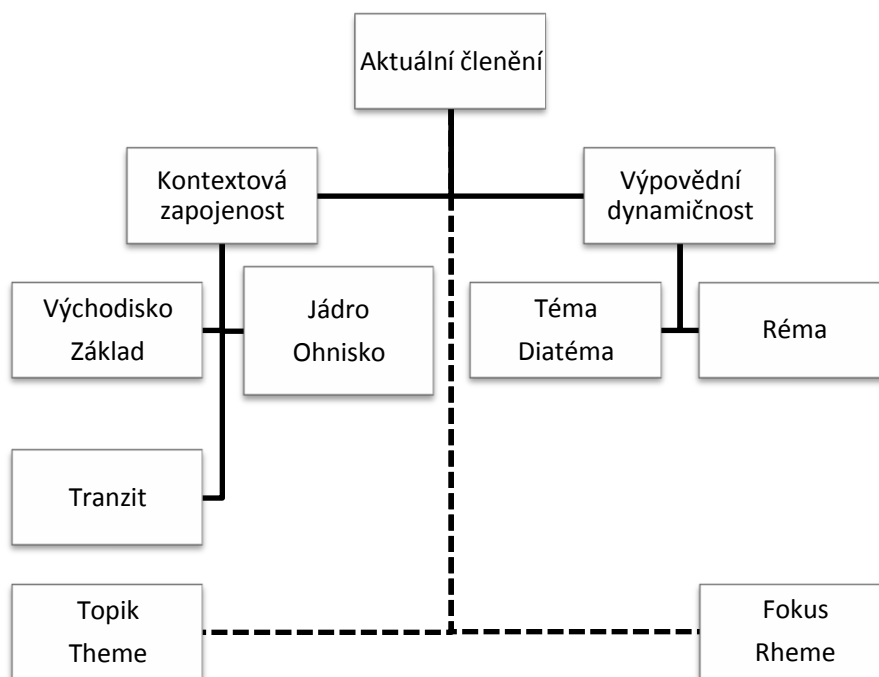
Aktuální členění (AČV) je jednou z perspektiv, kterou při nahlížení na koreferenci budeme v této práci uplatňovat. V této práci využíváme několika různých teoretických přístupů. Za základní považujeme důsledně aplikované rozlišování kontextové zapojenosti a výpovědní dynamičnosti po vzoru *topic-focus articulation* (TFA) a přístup Viléma Mathesia, který je zakladatelem českého zkoumání aktuálního členění, zejména jeho vymezení termínů základ, východisko a jádro, Sgallův přístup k definici ohniska a Firbasovo vymezení termínů výpovědní dynamičnosti (téma, réma). Kromě toho pracujeme s Danešovým pojetím tematických posloupností, které navazuje na Mathesia a Firbase.

Podle Mathesia (1947) je aktuální členění způsobem začlenění výpovědi do věcné souvislosti. Jde tedy o vztah uspořádání výpovědi z hlediska sémantického. Podle Daneše aktuální členění představuje především zachycení vzorců výstavby textu: „Ve skutečnosti dokonce i povrchní pozorování textů ukazuje, že výběr a rozdělení témat v textu ukazují určitý vzorec; toto tvrzení také odpovídá intuitivním očekáváním, že průběh prezentace předmětu komunikace musí nutně být řízen nějakými pravidelnostmi, musí být produktem vzorce.“ (Daneš 1974a: 109)<sup>5</sup> Daneš tedy vychází z předpokladu, že každá výstavba textu se řídí pravidly, která lze prostřednictvím aktuálního členění definovat, určit zákonitosti vztahující se k pravidelnostem výstavby textu. S tímto pojetím se ztotožňujeme, neboť vzorce výstavby textu lze podle našeho názoru snadno potvrdit empiricky, mimo jiné prostřednictvím zmapování principů výstavby koreferenčních řetězců, které je cílem této práce.

Stejným způsobem, jakým jsme zachytili termíny týkající se koreference, se v následujícím schématu B pokusíme zachytit i termíny týkající se aktuálního členění. Na základě hlediska kontextové zapojenosti rozlišujeme východisko, základ a jádro, ohnisko. Kromě toho bývá někdy vyčleňován zvlášť tzv. tranzit (viz dále). Hledisko výpovědní dynamičnosti člení výpověď na část tematickou a část rematickou. Je zřejmé, že kritéria kontextové zapojenosti a výpovědní dynamičnosti nelze striktně oddělovat, jsou to

<sup>5</sup> „In fact, even a superficial observation of texts shows that the choice and distribution of themes in the text reveal a certain patterning; this statement also corresponds to our intuitive expectations that the progression of the presentation of subject-matter must necessarily be governed by some regularities, must be patterned.“

perspektivy, které jsou propojeny, nicméně jejich rozlišování je podle nás účelné (viz dále). Členění na topik a fokus naznačené přerušovanou čarou, je spojením těchto dvou perspektiv, kdy termín topik odkazuje k částem výpovědi kontextově zapojeným a nejméně dynamickým, zatímco termín fokus odkazuje k částem výpovědi kontextově nezapojeným a z hlediska výpovědní dynamičnosti nejvíce dynamickým.



**Schéma B: aktuální členění**

V rámci textové lingvistiky jsou terminologické přístupy k aktuálnímu členění značně rozrůzněné, zejména ve srovnání přístupů jednotlivých autorů, ale i při srovnávání české tradice danešovské a mathesiovské s tradicí zahraniční, především americkou. Pokusím se proto zde odlišit termíny užívané nejčastěji a vymezit jejich vzájemný vztah alespoň v té míře, na níž panuje shoda. Snaha spojit anglické a původní české termíny je již předem odsouzena k neúspěchu, neboť lze jen těžko vzájemně se jen drobně odlišující české termíny spojit s jediným termínem anglickým, který při překladu bývá užíván pro více českých pojmů. Anglický termín *theme* bývá například používán pro označení základu i východiska výpovědi, termín *rheme* se překrývá s termínem *focus*, oba mohou označovat réma, fokus i ohnisko.

Termín aktuální členění zná česká lingvistika od Mathesia<sup>6</sup>, který ho postavil do opozice ke členění větnému, které nazýval formálním (viz výše). V literatuře bývají kromě původního Mathesiova termínu užívána i další označení, např. *functional sentence perspective* (FSP, funkční větná perspektiva, např. ÚJČ 1970, Daneš 1974, Firbas 1957, 1992), *information structure* (IS, informační struktura, např. Lambrecht 1994, Erteschik-Shir 2007, Mereu 2009) a již zmíněné *topic-focus articulation* (TFA, doslova artikulace topik a fokus, ale překládá se jako aktuální členění, např. Sgall 1998, Dvořáčková 2005), přestože nejsou plně synonymní.

Dvěma základními kritérii, která se při vymezování aktuálního členění v rámci *topic-focus articulation* uplatňují, jsou kontextová zapojenost a výpovědní dynamičnost. Tato distinkce se nám jeví jako odůvodněná, a proto s ní také pracujeme. Zatímco kontextová zapojenost implikuje dvě diskrétní hodnoty, je tedy binární, výpovědní dynamičnost ze své podstaty odkazuje ke skalárnímu uspořádání. Přínosnost rozlišování těchto dvou kritérií je zdůvodněna i M. A. K. Hallidayem, který ve svém článku *Notes on transitivity and theme in English* z roku 1967 rozlišuje „dvě simultánní struktury textu: (1) 'informační důraz' (dané – nové), a (2) 'thematizace' (T – R)“ (Daneš 1974a: 107)<sup>7</sup>.

Nyní definujeme termíny, s nimiž v rámci zpracování aktuálního členění v této práci pracujeme, tedy kontextovou zapojenost a výpovědní dynamičnost.

### 2.2.1 Kontextová zapojenost

Kritérium kontextové zapojenosti člení částí promluvy na kontextově nezapojené (jádro, ohnisko) a kontextově zapojené (základ, východisko). Kritériem zařazení částí výpovědi do kontextově zapojené nebo kontextově nezapojené skupiny je novost informace, která (podle Hallidaye) nespočívá v tom, že informace nebyla v předchozím kontextu zmíněna, ale v tom, že ji mluvčí prezentuje jako informaci, která není odvoditelná z předchozího kontextu: „Halliday upozornil na to, že nová informace není 'nová' v tom smyslu, že

<sup>6</sup> Mathesiova stať *O tak zvaném aktuálním členění větném* vyšla poprvé ve Slově a slovesnosti v roce 1939 (ročník 5, str. 171–174), v seznamu literatury a při citování je odkazováno k jejímu vydání v souboru Mathesiových statí *Čeština a obecný jazykozpyt* z roku 1947.

<sup>7</sup> „two simultaneous structures of text: (1) 'information focus' (given – new), and (2) 'thematization' (T – R)“

‘nebyla zmíněna před tím, přestože to často tento případ je, ale v tom smyslu, že mluvčí ji prezentuje jako nezískatelnou z předchozího kontextu’.“ (Daneš 1974a: 110)<sup>8</sup>

Perspektiva kontextové zapojenosti je klíčová i z toho důvodu, že je základním kritériem pro rozlišení hodnot atributu tfa v PDT (viz dále). Proto ji zde představíme prostřednictvím definování termínů základ, východisko, jádro a ohnisko.

Při určování toho, zda část výpovědi patří ke skupině kontextově zapojených prvků, nebo patří ke skupině prvků kontextově nezapojených, pracujeme s Mathesiovými (1947) termíny základ, východisko (východiště) a jádro, které každý odkazují k trochu jiné perspektivě. Termín ohnisko byl zaveden v rámci funkčního generativního popisu a vychází z aplikace stromové hierarchické struktury. Při jeho vymezení vycházím z pojetí představeného Petrem Sgallem (např. 1972). Nejprve definujeme termíny základ a východisko.

#### 2.2.1.1 Základ

Základem výpovědi podle Mathesia je to, o čem se mluví. K vymezení základu je oproti vymezení východiska uplatňováno kritérium obsahového členění („o čem se mluví“ odkazuje k sémantice).

#### 2.2.1.2 Východisko

Mathesius vymezuje východisko prostřednictvím perspektivy textové – na základě vymezení kontextově známé informace: „charakteristika východiska [...] byla založena na členění kontextovém (co je známé, co je nasnadě)“ (ESČ 2002: 562).

Důsledné rozlišování základu a východiska nenacházíme ani v pozdějších Mathesiových pracích (Firbas 1996), dnes základ a východisko splývají: „[...] dnes je situace taková, že [...] se obě hlediska promítají do jednotného výsledku, jímž je téma, tematická jednotka, popřípadě tematická sféra věty (v angličtině *theme* nebo *topic*)“ (ESČ 2002: 534).

Určení známé (dané) informace je samo o sobě problematické a ovlivněné několika dílčími faktory. Prvním z nich je fakt, že podstatou, na jejímž základě je určována znalost,

<sup>8</sup> „Halliday has pointed out that the new piece of information is ‘new’ not in the sense that ‘it has not been previously mentioned although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse’.“



resp. neznalost informace z kontextu, je hodnocení prováděné mluvčím a adresátem. Z toho samozřejmě vyplývají problémy související s graduálností tohoto hodnocení, se subjektivním hodnocením známosti jednotlivých informací. Daneš k tomu také poznamenává, že „danost závisí na délce předchozího textu, vzhledem k němuž je hodnocení prováděno“ (Daneš 1974a: 109)<sup>9</sup>.

Následující část textu bude věnována definici jádra a ohniska výpovědi, tedy prvkům spadajícím do sféry kontextově nezapojených výrazů. Jádro je Mathesiův termín, stejně jako předcházející základ a východisko (východiště), ohnisko je termín, s nímž pracuje FGP (viz dále).

### 2.2.1.3 Jádro

Jádrová část výpovědi stojí v opozici k základu a východisku popsaným výše. Jádro je charakterizováno opačnými rysy než kontextově zapojené části výpovědi, tj. je netematické a představuje „všechny větné složky, které nebyly součástí východiska, případně základu výpovědi“ (EŠČ 2002: 191).

Jádro chápeme jako složky výpovědi čistě kontextově nezapojené, ale i složky tranzitní, přičemž vlastní tranzit je „[v] teorii FSP jedna z komunikačních jednotek; tvoří pomyslný předěl mezi tematickou a netematickou sférou věty [...]. Vlastní tranzit je tedy jednotka, která v distribučním poli VD nese vyšší relativní stupeň VD než kterákoli jiná tranzitní či rematická jednotka.“ (EŠČ 2002: 496) Tranzit není v tradičním slova smyslu přesným protějškem kontextově zapojeným složkám výpovědi, neboť kritérium kontextové zapojenosti zařazuje složky výpovědi pouze do dvou skupin, do skupiny prvků kontextově zapojených a do skupiny prvků kontextově nezapojených. V takovéto klasifikaci nejsou složky tranzitní nijak specifikovány.

Z uvedeného vymezení základu, východiska a jádra je zároveň zřejmé, že v Mathesiově pojetí je důraz kladen na vyčlenění tematické složky výpovědi oproti ostatním částem věty. Důsledkem tohoto přístupu je spojování prvků tranzitních kontextově nezapojených, tedy

---

<sup>9</sup> „'Givenness' depends on the length of the proportion of preceding text in relation to which the evaluation is being carried out.“

vytváření prvků jádrových, které ovšem lze tolerovat vzhledem k nevyhraněnosti a sémantické nevýraznosti tranzitních prvků.

#### 2.2.1.4 Ohnisko

Ohniskem se rozumí ta část věty, která v teorii FGP (ÚJČ 1970) není součástí základu věty. Jako ohnisko bývá tedy označována ta část výpovědi, která není kontextově zapojená, není součástí tzv. horního podstromu věty a zároveň není prvkům kontextově zapojeným ani částem horního podstromu podřízená. V podstatě se tedy jedná o prvky větné struktury, které nevycházejí z valence slovesa, neobsazují valenční pozice a ani nejsou těmto valenčním pozicím podřízeny. V tomto smyslu je ohnisko specifičtější a z jiné perspektivy (totiž syntakticky namísto sémanticky) vymezenou částí výpovědi. Vzhledem k této velké odlišnosti s termínem ohnisko nepracujeme, uvádíme ho jen pro úplnost.

Podrobnější charakteristiku ohniska podává například Petr Sgall ve svém článku *Základ a jádro, nebo presupozice a ohnisko?* (1972), v němž mimo jiné shrnuje, že „umístění hranice ohniska je dáno kontextem“ (Sgall 1972: 163).

Vztah mezi jádrem a ohniskem není jednoznačný. Jádro je charakterizováno pomocí negativního vymezení vůči východisku (všechno, co není východiskem, je jádrem), s valenční strukturou se nepracuje. Oproti tomu v případě ohniska je důraz kladen na vazby mezi slovesem a jeho valenčními doplněními, které určují, zda je určitá část výpovědi ohniskem, či nikoli.

#### 2.2.2 Výpovědní dynamičnost

Definicí termínů základ, východisko, jádro a ohnisko jsme vyčerpali termíny spadající do sféry kontextové zapojenosti, nyní se přesuneme k výpovědní dynamičnosti. Druhým kritériem vedle kontextové zapojenosti uplatňovaným v přístupu k jazykovým strukturám z hlediska aktuálního členění je výpovědní dynamičnost (VD). Stejně jako v případě principu kontextové zapojenosti aplikujeme kritérium VD na jazykové struktury a jejich části jsou členěny na ty s nízkou a na ty s vysokou VD. VD je kritérium kopírující novost, přínosnost informace nesené danou částí výpovědi. Čím novější informace (ve smyslu její neodvoditelnosti z kontextu), tím vyšší stupeň VD je v dané části.

Distinkce založená na kritériu výpovědního dynamismu (*communicative dynamism*) byla zavedena Janem Firbasem (např. 1986)<sup>10</sup>, který VD charakterizuje jako „míru, s jakou větný prvek přispívá k rozvoji komunikace“ (Daneš 1974a: 106n.)<sup>11</sup>, VD je „ve Firbasově pojetí [...] hnací síla, kterou v okamžiku promluvy přispívají jednotlivé složky výpovědi (nositelé výpovědní dynamičnosti) k rozvoji sdělení“ (ESČ 2002: 120).

Výrazný rozdíl mezi kontextovou zapojeností a výpovědní dynamičností nacházíme ve faktu, že v případě kontextové zapojenosti jsou jejím nositelem vždy celá slova, respektive fráze, kdežto v případě výpovědní dynamičnosti můžeme rozlišovat dynamičnost i některých morfémů ve slově. V rámci výkladu týkajícího se VD v ESČ (2002: 120) je analýze podrobena jednoduché syntagma *udělala kotrmelec*, které z hlediska výpovědní dynamičnosti člení Aleš Svoboda (autor hesla v ESČ) takto (řazeno od nejnižšího stupně VD): morfém *-a* < morfém *-l-* < nociální složka slovesa *uděla-* < *kotrmelec*. Takováto interpretace výpovědní dynamičnosti samozřejmě vychází z předpokladu neutrální situace (běžné okolnosti, výpověď bez citového zabarvení, ...) a výpověď *Ten kotrmelec udělala ona*. bude mít z hlediska struktury VD jiné uspořádání.

V rámci výpovědní dynamičnosti nyní představíme dva klíčové termíny. Těmi jsou téma a réma.

#### 2.2.2.1 Téma

K tematické části výpovědi jsme odkazovali již v části týkající se kontextové zapojenosti. Nicméně na tomto místě se budeme věnovat tématu samotnému. Tématem bylo „původně to, o čem se mluví, v protikladu k tomu, co se o tom říká“ (ESČ 2002: 479). Téma je jednou z krajních položek škály VD, konkrétně tou nejméně dynamickou, jinými slovy má téma nejnižší stupeň výpovědní dynamičnosti ze všech složek výpovědi. Téma můžeme charakterizovat i poněkud vágněji, a to ne z hlediska struktury vyjádření ve smyslu „o čem“ a „co“, ale z hlediska známosti jako tu část výpovědi, která je z kontextu známá nebo odvoditelná. Tímto způsobem vymezení tématu je v našem pojetí propojena výpovědní dynamičnost s kontextovou zapojeností.

<sup>10</sup> J. Firbas je autorem termínů téma a réma.

<sup>11</sup> „the extent to which the sentence element contributes to the development of the communication“

## 2.2.2.2 Réma

Réma stojí v opozici k tématu a je vymezeno vůči němu tak, že téma, jak bylo řečeno výše, je to, o čem se mluví, zatímco réma je to, co se o tom říká. Takováto definice poslouží pro základní rozlišení, nicméně není příliš exaktní. Pro přesnější určení hranic rématu slouží v jazyce například tzv. rematizátory (*taky, však, jenom, právě*), které signalizují hranici mírně a silně dynamických prvků výpovědi. Kromě nich vychází určení rématu zejména z komunikační situace. Rématem je ta část výpovědi, která má nejvyšší stupeň výpovědní dynamičnosti, která je nejnovější, nejvíce informativní a nejméně známá. Opět je zřejmé, že se při vymezování rématu nevyhneme smíšení kritérií dynamičnosti a kontextové zapojenosti.

Téma ani réma tedy nejsou absolutními kategoriemi v tom smyslu, že v jejich rámci jsou tradičně vyčleňovány centrální a periferní prvky. Za centrální prvky kategorií tématu a rématu můžeme považovat vlastní téma a vlastní réma, která v rámci hierarchie vytvořené na základě kritéria VD považujeme za nejvíce (v případě rématu) a nejméně (v případě tématu) dynamické prvky nejen v rámci výpovědi jako celku, ale také v rámci celého tématu, resp. rématu.

Problematicku aktuálního členění a jeho terminologie zachycuje následující modelový příklad:<sup>12</sup> V1: <[Petr<sub>VT</sub>]<sub>T</sub> [poslala]<sup>V</sup> <babička dopis<sub>VR</sub>><sup>J</sup>]<sub>R</sub>. V2: <[Dopis/Ten<sub>VT</sub>]<sub>T</sub>><sup>V</sup> <však [nikdy nedorazil<sub>VR</sub>]<sub>R</sub>><sup>J</sup>.

Pokud se budeme držet původní mathesiovské terminologie představené výše, za východisko výpovědi V1 označíme její podmět a predikát, za její jádro (J) pak pravovalenční doplnění predikátu v dativu a akuzativu. Je zřejmé, že hranice tématu (T) a rématu (R) a východiska (V) a jádra (J) se nepřekrývají, zřetelně je vidět například odlišnost v rámci hranic jádra a rématu. V rámci V2 se nachází rematizátor *však*, který signalizuje začátek rematické části výpovědi. V naší interpretaci je již součástí rématu, v jiném pojetí by mohl být vymezen mimo téma i mimo réma a plnit funkci konektoru, signalizovat vztah mezi V1 a V2. V rámci distinkce T – R vymezujeme i vlastní téma (VT) a vlastní réma (VR). Tyto jednotky jsou označeny pouze na úrovni slov, nikoli v jejich

<sup>12</sup> Hledisko kontextové zapojenosti je zachyceno horními indexy a špičatými závorkami, hledisko výpovědní dynamičnosti dolními indexy a hranatými závorkami.

rámci, a jsou značeny přímo za slovem, které zastává pozici vlastního T, resp. R. Slovesný predikát řadíme již do části rematické z toho důvodu, že tranzit, tedy jednotky přechodné, tj. ani tematické, ani rematické, nevymezujeme a z hlediska výpovědní dynamičnosti je sloveso spíše dynamičtější složkou vyjádření a má tedy blíže k rématu.

Na interpretaci hranic jednotlivých celků může mít vliv velké množství faktorů, zejména intonace, konkrétní komunikační situace a perspektiva z komunikační situace vycházející.<sup>13</sup> Nicméně uvedený příklad je uměle vytvořený a čistě modelový a značení hranic jednotlivých celků a interpretace jejich kontextové zapojenosti a výpovědní dynamičnosti slouží pouze účelům demonstrativním.

Vymezením tématu a rématu jsme vyčerpali skupinu termínů týkajících se výpovědní dynamičnosti. Nyní se vymežíme termíny, vycházející ze zahraničních lingvistických teorií, topik a fokus.

### 2.2.3 *Topik*

Hajičová definuje topik jako hlavní téma diskurzu (což bychom v pojetí danešovském pravděpodobně označili za hypertéma) a dodává, že „je možné stanovit téma (či přesněji řečeno témata) diskurzu jako soubor prvků, které se po určitou delší dobu udržely na nejvyšších stupních aktivovanosti“ (Hajičová 1999: 51). Toto pojetí vychází, stejně jako celá perspektiva topiku (fokusu), z přístupu, který je vlastní i centeringové teorii (viz dále).

Vzhledem k tomu, že ve zkoumání navazování koreferenčních řetězců vycházíme ze syntaktické teorie, která má původ i terminologii ukotvenou v anglicky mluvícím prostředí, a vzhledem k velkému množství českých termínů označujících kontextově zapojenou část výpovědi, které by mohly svádět k příliš zúženému pohledu na problematiku navazování koreferenčních vztahů v textu, jsme se rozhodli užívat termínu *topik* namísto základu, východiska, jádra a ohniska, ale i tématu.

<sup>13</sup> Jedním z nezanedbatelných faktorů je i to, že dichotomie není úplná: zatímco v rámci tematické části výpovědi dochází k rozlišení tématu (o čem se mluví) a východiska/základu (známá informace), v rámci rematické jsou réma i jádro vymezeny shodně (co se o tématu říká).

#### 2.2.4 Fokus

Fokus bývá definován negativně, tj. ve vztahu k topik. Jaroslav Peregrin říká, že „všechny kontextově zapojené prvky závislé na predikátu společně s prvky na nich závislými a s predikátem samotným, pokud je kontextově zapojený, utváří *topik* věty; zbytek věty tvoří *fokus*“ (Peregrin 1995: 50).<sup>14</sup>

Stejně meta-terminologicky jako topik užíváme termín fokus, slouží tedy k označení těch částí výpovědí, které jsou kontextově nezapojené a z hlediska výpovědní dynamičnosti obsazují její nejvyšší stupně v rámci diskurzu.

Termínu topik tedy užíváme jako souhrnného „metatermínu“, který vychází nejen z kritéria kontextové zapojenosti, ale zahrnuje i výpovědní dynamičnost, čili v případě nejméně dynamické části výpovědi téma. Na stejném principu funguje i „metatermín“ fokus. Například ve výpovědi *Petr vyhrál závod*, představuje výraz *Petr*, který bychom označili jako topik, prvek kontextově zapojený a nejméně dynamický, zatímco výraz *závod*, který bychom označili jako fokus, prvek kontextově nezapojený a zároveň prvek nejvíce dynamický z hlediska výpovědní dynamičnosti.

Dosud jsme se věnovali problematice koreference a aktuálního členění. Definovali jsme termíny týkající se koreference, i termíny týkající se aktuálního členění. Problematika definice jednotlivých termínů týkajících se koreference a aktuálního členění, s nimiž v jeho rámci operujeme, je tímto završena. Nyní přejdeme k různým přístupům, které jsou vzhledem k aktuálnímu členění uplatňovány.

### 2.3 Větný člen

Nyní se přesuneme k druhému hledisku, které můžeme při analýze struktury výpovědi aplikovat, a tím je přiřazování větněčlenské platnosti jednotlivým výrazům. Toto kritérium není z našeho pohledu problematické. Větný člen je obecně definován jako „syntakticky relevantní prvek/element větné struktury, ke kterému se dospívá na základě syntaktické analýzy“ (ESČ 2002: 100). Tato definice je záměrně vágní, aby byla schopna obsáhnout

<sup>14</sup> „all the contextually bound items depending on the main verb together with all that depends on them and together with the main verb if this is contextually bound, constitute the *topic* of the sentence; the rest of the sentence constitutes the *focus*“

všechny typy větných členů. Nyní přejdeme ke dvěma větným členům, které jsou pro naši analýzu klíčové, tj. k podmětu a předmětu.

### 2.3.1 Podmět

Podmět bývá spolu s přísudkem považován za nejdůležitější větný člen, protože spolu dohromady podmět a přísudek vytvářejí základní větnou strukturu, predikační dvojici.

Podmět je definován právě prostřednictvím svého vztahu k přísudku jako „větný člen, jemuž se přísudkem přisuzuje nějaký příznak (děj, stav a jejich změny)“ (ESČ 2002: 316). Zároveň má podmět další charakteristické rysy: je syntakticky nezávislý, bývá v neutrálním vyjádření nejexponovanějším slovním vyjádřením a z toho důvodu také nejvýraznějším centrem pozornosti, bývá často vyjádřen zájmenem nebo bývá povrchově nevyjádřen, podle některých teorií může být vyjádřen i jinými pády než nominativem (např. genitiv, lokál) apod.

Pro tuto práci jsou klíčovými charakteristikami podmětu především jeho syntaktická nezávislost v rámci struktury výpovědi, možnost jeho upozadění apod. (primární a sekundární hierarchizace větné propozice) a možnost postavení podmětu v antecedentu koreferenčního řetězce.

### 2.3.2 Předmět

Oproti podmětu je předmět syntakticky závislý ve smyslu formálním (pád předmětu je určen slovesnou valencí, „v tradiční gramatice [je předmět] větný člen, který rozvíjí sloveso nebo adjektivum (jakékoli větněčlenská platnost) a jehož tvar je tímto slovesem nebo adjektivem řízen“ (ESČ 2002: 350).

V textu se předmět vyskytuje zejména v případech, kdy sloveso otevírá dvě a více valenčních pozic, z nichž typicky jedna náleží podmětu a druhá předmětu. Případné další pozice jsou obsazovány různými dalšími větnými členy nebo dalšími reprezentanty již vyjádřených větných členů. Možnost reduplikace pozice předmětu (tedy ne stejné entity v postavení předmětu, ale pozice předmětu, které nabízejí možnost obsazení různými entitami plnícími funkci předmětu) je dalším charakteristickým rysem, který odlišuje podmět a předmět – zatímco podmět je vždy jen jeden (a může být rozšiřován znásobením v rámci jedné podmětové pozice, jako např. *Petr a Pavel koupili auto.* nebo *Petr s Pavlem*

*koupili auto.*), předmětů může být v jedné struktuře více (např. *Petr koupil mamince nové auto.*)

Předměty bývají většinou vyjádřeny jmennou frází, oproti podmětu, jehož prototypickou realizací v češtině je jeho nevyjádření na povrchu. Rozlišování přímého a nepřímého předmětu nemá vliv na navazování vztahů v textu, proto se jím nebudeme zabývat. Bližší pojednání o podmětu a předmětu podáme v kapitole věnované Ludmile Zimové a jejím zkoumáním (3.3)

Ve druhé kapitole jsme představili termíny, s nimiž budeme nadále pracovat. V sekci věnované koreferenci to byly termíny *antecedent*, *anafor* a *zájmeno třetí osoby*, v sekci věnované aktuálnímu členění termíny *kontextová zapojenost*, *základ*, *východisko*, *jádro*, *ohnisko*, *výpovědní dynamičnost*, *téma*, *réma*, *topik* a *fokus*. Z větných členů jsme se zaměřili na podmět a předmět. V následujících kapitolách budou představeny dosavadní přístupy k navazování koreferenčních vztahů využívající představenou terminologii a korpusový a dotazníkový výzkum věnovaný stejné problematice.



### 3 Dosavadní přístupy

V této části práce se budeme věnovat dosavadním přístupům k aktuálnímu členění a větněčlenské platnosti z hlediska rysů, které jsou klíčové pro utváření koreferenčních řetězců. Jen krátce zmíníme Mathesia, o jehož přínosu k terminologii aktuálního členění již byla řeč, podrobněji popíšeme koncept Danešových tematických posloupností, aby bylo zřejmé, jaké způsoby výstavby textu jsou v češtině doloženy a zároveň aby bylo možné propojit českou tradici zkoumání výstavby textu reprezentovanou právě Danešem s teorií zahraniční reprezentovanou centeringovou teorií.

Přístup zachycující jednotlivé způsoby vyjadřování větných členů v textu, který spolu s teorií centeringu slouží jako základ pro hypotézy o navazování na jednotlivé větné členy, reprezentuje práce Ludmily Zimové *Vyjadřování větných členů v textu* (1994), která byla vybrána proto, že přistupuje k vyjadřování větných členů v textu z hlediska hypersyntaktického, tedy z hlediska výstavby textu. Zimová rozebírá i různé způsoby výstavby textu odvíjející se od formy, v níž je větný člen realizován (elize, pronominalizace, plná fráze). Práce všech těchto tří badatelů jsou z dnešního pohledu relativně staré (nejnovější je monografie Ludmily Zimové z roku 1994), mimo jiné proto jsme se rozhodli pro zkoumání aktuálního členění použít i teorii zahraniční, v českém prostředí novou, tj. teorii centeringu. Centering navíc poskytuje predikce o četnosti navazování vztahů v textu z hlediska větněčlenské platnosti i kontextové zapojenosti.

#### 3.1 Vilém Mathesius

Vilému Mathesiovi (1882–1945) jsme se věnovali již v předešlé části diplomové práce, v níž jsme definovali užívané termíny. Z toho důvodu ho jakožto zakladatele české aktuálněčlenské terminologie na tomto místě jen připomínáme.

#### 3.2 František Daneš

František Daneš (1919–) jakožto autor kanonického textu o tematických posloupnostech představil pět typů budování textu,<sup>15</sup> z nichž za základní považujeme průběžné téma

<sup>15</sup> Tematické posloupnosti jsou ve své podstatě také navazováním vztahů v textu, neboť jednotlivé typy vymezují vazby mezi částmi aktuální a předchozí výpovědi. Přesto se o nich v tomto smyslu nevyjadřujeme, jako navazování vztahů v textu označujeme pouze vztahy koreferenční, i když se někdy mohou se vztahy tématu (a rématu) aktuální výpovědi s tématem/rématem výpovědi předchozí překrývat.

a tematizaci rématu představené dále. Aby bylo vůbec možné tematické posloupnosti obecně zkoumat, musíme nejprve vymežit části textu, mezi nimiž jsou vztahy detekovány. V tomto bodě navazuje Daneš na Mathesia (základ, východisko, jádro), proto bychom rádi alespoň částečně usouvztažnili Mathesiovu klasifikaci s Danešovou perspektivou, která se vůči ní vymezuje tak, že na přesné rozlišování dvou aspektů mathesiovského členění, tedy členění tématu – rématu a členění východiska – jádra, rezignuje, „poněvadž nám v našem článku nejde o řešení vlastní problematiky aktuálního členění“ (Daneš 1968: 126). Daneš tedy dělení na téma – réma užívá jako metatermín (stejně jako v našem případě užíváme topik a fokus) k odkazování na široce vymezenou tematickou a rematickou část výpovědi.

Stejně jako Mathesius považuje Daneš za důležité rozlišovat funkci textové výstavby, kterou označuje jako funkci výpovědní (v Mathesiově pojetí této funkci odpovídá hledisko aktuálněčlenské), a funkci větněčlenskou, kterou nazývá funkcí větně konstrukční (v Mathesiově pojetí se jedná o hledisko formálního členění). Zároveň s tímto dělením chápe Daneš větu a výpověď jako „útvár jediný, v němž rozeznáváme různé jeho aspekty a složky, různé funkční roviny jeho stavby“ (ibid: 128).

Téma a réma pro Daneše představují „dvě dílčí komplementární funkce výpovědi, kterých nabývají různé části věty, když je řazena do kontextu jako jeho součást s platností výpovědi, přesně řečeno, jde o funkce významů větných částí“ (ibid: 127), přičemž „v klidném, věcném sdělení [...] umísťujeme T na počátek výpovědi a R pak na její konec“ (Daneš 1999: 109). Role kontextu je nezastupitelná, stejně jako přítomnost obou komplementárních složek funkcí výpovědi: každá výpověď má tematickou i rematickou část, které jsou vyčleněny na základě široce chápaného kontextu. Z příkladů, které Daneš uvádí, je zřejmé, že nepracuje s žádnou formou tranzitu, slovesné predikáty řadí do části rematické, předpokládá tedy, že predikátové sloveso je součástí toho, co se o věci říká, ne toho, o čem se mluví. V případě podřadicího souvětí to platí také, např. *[Abychom takto získané surové železo přeměnili na ocel dobré jakosti]<sub>T</sub>, [musíme z něho odstranit škodlivé příměsi...]<sub>R</sub>*. Jednotlivé věty tvořící souvětnou strukturu jsou analyzovány zvlášť: *[Zde]<sub>T</sub> [probíhá oxidování ostatních příměsí železa]<sub>R</sub> a [surové železo]<sub>T</sub> [se proměňuje v ocel]<sub>R</sub>*. S tímto pojetím se ztotožňujeme. V případě koordinace ne na úrovni větné, ale v rámci úrovně větněčlenské mluví Daneš (1968) o několikanásobném rématu (viz dále)

a o několikanásobném tématu (o možnostech vyjádření tématu také viz dále v rámci tzv. komplexních výpovědí).

Kontext je u Daneše charakterizován prostřednictvím otázkového testu:

*„Zdá se však, že lze najít kritérium pro zjištění všech různých typů kontextů, v nichž se může daná věta objevit vždy v jedné ze svých různých výpovědních modifikací: typem kontextu v tomto smyslu bychom pak rozuměli právě tu podmnožinu množiny všech možných kontextů dané věty, v níž se tato věta vyskytuje v jedné ze svých modifikací. Jinými slovy: jde o kritérium umožňující rozhodnout o kterémkoli kontextu, a) zda v něm lze dané věty užít, b) v kladném případě pak, ke které výpovědní modifikaci (popř. ke kterým výpovědním modifikacím) dané věty patří. Postupem, který splnění tohoto úkolu umožňuje, je vyhledávání otázek.“* (Daneš 1968: 127)

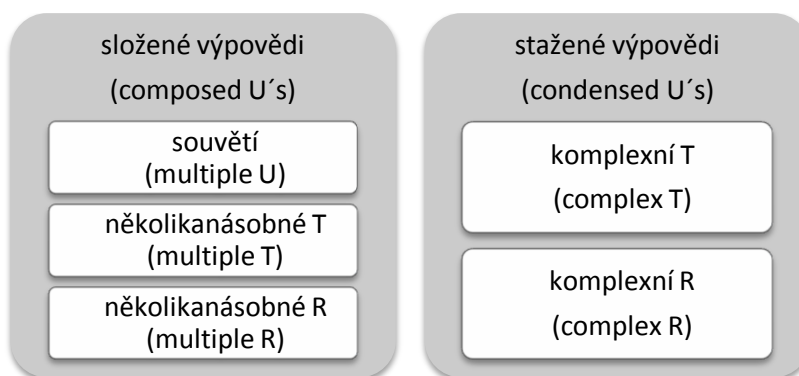
Daneš tedy ke každé větě oznamovací, kterou podle něj lze považovat za odpověď, hledá otázku, na niž by byla vhodnou a adekvátní odpovědí. Zároveň z toho vyplývá, že „různé jednotlivé otázky [...] vymezují (reprezentují) všechny různé typy kontextu této věty (neboť různým výpovědním modifikacím dané věty odpovídá právě některá z množiny otázek přiřazených k oné větě)“ (ibid:128). Otázkový test je omezen na rematickou část výpovědi, neboť „na téma se ptát nedovedeme, leda paušální, zcela obecnou otázkou ‘O čem se v této výpovědi mluví (T) a co se o tom říká (R)?’“ (ibid: 129).<sup>16</sup> Tematická část výpovědi totiž slouží ke stavbě otázky, kdežto rematická utváří odpověď na ni. Toto rozdělení vyplývá z toho, že téma je částí výpovědi známou, kdežto réma je informací novou, obohacující, tudíž pokud pokládáme výpověď za odpověď na otázku, nemůže k jejímu sestavení sloužit ta část výpovědi, která je známá až ve chvíli jejího zodpovězení (a naopak). Podle Daneše jediný způsob, jak se „zeptat“ na téma, je parafráze: „témata výpovědi lze zpravidla parafrázovat pomocí fráze ‘pokud jde o...’“.

Ustavení pěti typů tematických posloupností (TP) vychází z toho, že „si můžeme vzhledem k souvislé promluvě (textu) položit otázku, odkud se témata jednotlivých výpovědí berou, jaký je vzájemný vztah témat výpovědi patřících k témuž odstavci a jaký je jejich vztah k nadřazenému ‘hypertématu’ tohoto odstavce (popř. k nadřazeným tématům vyšších tematických celků)“ (ibid: 126). Pokud se pokoušíme zodpovědět tyto otázky,

<sup>16</sup>      Není ovšem tak docela pravda, že při určování tématu nemá otázkový test co poskytnout, neboť s vymezením rématu dostáváme i indicie toho, kde jsou potenciálně tematické části výpovědi.

docházíme k následujícím pěti základním typům výstavby promluvy (ve smyslu textu, několika souvislých výpovědí): typ s návaznou tematizací rématu, s průběžným tématem, s rozvíjením rozštěpeného rématu, s tematickým odvozováním a s tematickým skokem.<sup>17</sup> Tento pětičlenný výčet typů tematických posloupností není vyčerpávající, je to „stručný a neúplný výčet nejtypičtějších a nejběžnějších způsobů rozvíjení tematické linie“ (Daneš 1999: 111). Danešem ustavené tematické posloupnosti jsou podle něj základní reprezentací textové koherence (*text connexity*) – viz Daneš 1974a.

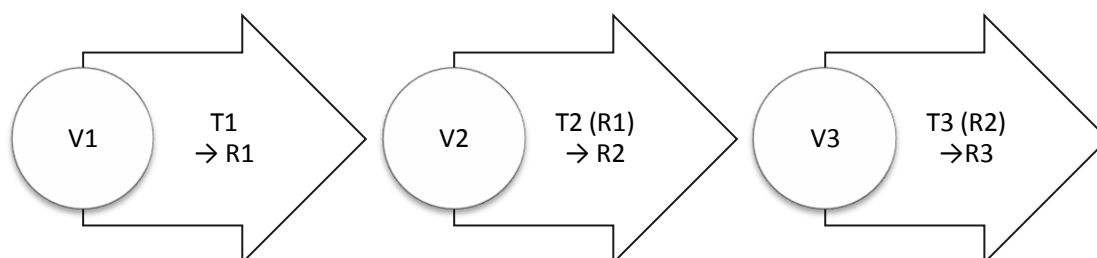
Napříč vyčleňováním těchto pěti základních typů Daneš vymezuje ještě tzv. složené výpovědi (*composed U's*), které jsou tří typů: výpovědi několikanásobné (*multiple U*), výpovědi s několikanásobným tématem (*multiple T*) a výpovědi s několikanásobným rématem (*multiple R*) a tzv. stažené výpovědi (*condensed U's*), které jsou dvou typů: výpovědi s komplexním tématem (*complex T*) a výpovědi s komplexním rématem (*komplex R*). Schematicky zachytíme typologii složených a stažených výpovědí v následujícím schématu C, ale dále se jim věnovat nebudeme, neboť nejsou pro tuto práci klíčové:



**Schéma C: složené a stažené výpovědi**

Nyní představíme postupně pět typů tematických posloupností tak, jak je uvádí Daneš (1968, 1974a, 1999 apod.).

<sup>17</sup> Příklady uváděné v této sekci jsou čerpány z Daneše (1968: 128n.), pokud není uvedeno jinak.

3.2.1 Návazná tematizace rématu<sup>18</sup>Schéma D: návazná tematizace rématu<sup>19</sup>

Prvním a podle Daneše (1968) nejčastějším typem tematických posloupností v textu je tematizace rématu. Jejím základním znakem je, že tématem postcedentní věty se stává réma věty antecedentní, jde tedy o postupné rozvíjení informace  $A \rightarrow B \rightarrow C \rightarrow D$ . Typicky je tohoto schématu užíváno „v prostém výkladu učebnicového typu, v jednoduchých popisech různých technologických a podobných postupů, v popise událostí, při referování atd.“ (Daneš 1968: 133). Příkladem takovéto výstavby tematické posloupnosti u Daneše je kontext *Dne 10. března sovětská vláda odevzdala vládám tří velmocí nótu o mírové smlouvě s Německem. K nótě byl přiložen návrh osnov mírové smlouvy. Tento sovětský návrh tvoří začátek nové etapy ve vývoji otázky mírové smlouvy.*, který přesně odpovídá uvedenému schématu.

<sup>18</sup> V rámci všech schémat uvádím z důvodu přehlednosti a úspory prostoru znázornění pouze tří po sobě bezprostředně následujících výpovědí s tím předpokladem, že se cyklicky opakuje vztah mezi V2 a V3.

<sup>19</sup> V tomto i následujících schématech představuje každá velká šipka jednu výpověď (její označení zahrnující pořadí je uvedeno v kroužku, např. V1, V2 apod.). T a R s příslušným očíslováním představují označení tématu a rématu příslušné výpovědi. Šipky mezi T a R naznačují lineární řazení výrazů tvořících T, resp. R, ve výpovědi. Závorky představují vztah mezi T výpovědi aktuální a předcházejícími T, resp. R. Např. T2 (R1) znamená, že tématem druhé výpovědi (T2) je réma výpovědi předcházející (R1).

## 3.2.2 Průběžné téma

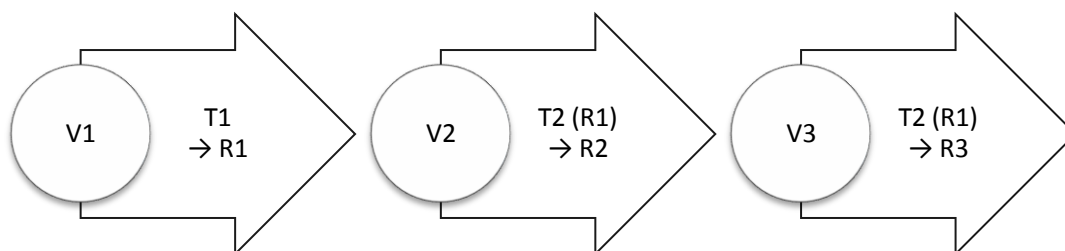


Schéma E: průběžné téma

Schéma tematické posloupnosti průběžného tématu je typické tím, že „zůstává téma, navozené např. rématem úvodní výpovědi, stále totéž (opakuje se) a přiřazují se k němu vždy nová rémata“ (Daneš 1968: 134). Zajímavým a na tomto schématu patrným jevem je fakt, že první výpověď je považována za úvod, nastolení hypertématu, až druhá a následující výpovědi jsou určující pro stanovení schématu tematické posloupnosti. Z toho důvodu je znázorněné schéma hodnoceno jako tematizace rématu. Existují samozřejmě i varianty, v nichž bychom vztah mezi V1 a V2 hodnotili ne jako tematizaci rématu, ale už na této úrovni bychom našli průběžné téma.

Tento typ tematické posloupnosti „je častý např. při popise jednotlivého jevu předmětu nebo výkladu o něm (tématem zůstává tento jev a uvádějí se jednotlivě jeho vlastnosti ap.)“ (ibid: 134) a příkladem je kontext *Mezi tato tzv. ovládací automatická zařízení patří např. většina automatických výrobních linek. Tyto linky a jim podobná zařízení vykonávají samočinně všechny operace nutné k zhotovení výrobku. Tato zařízení však nemohou kontrolovat průběh a výsledky své činnosti. Nejsou také schopna přizpůsobit se změnám vnějších podmínek nebo změnám svého vlastního vnitřního stavu.*

## 3.2.3 Rozvíjení rozštěpeného rématu

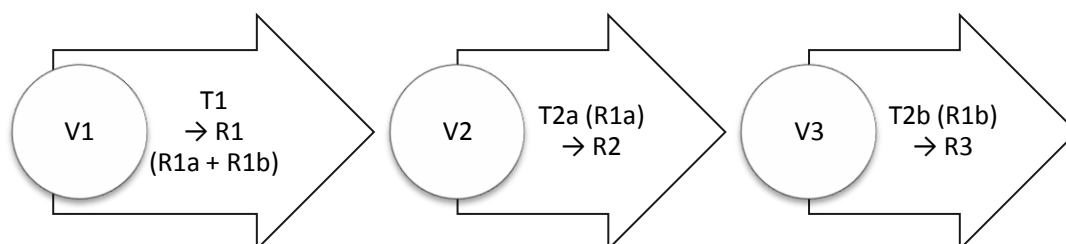


Schéma F: rozvíjení rozštěpeného rématu

Tematická posloupnost rozvíjení rozštěpeného rématu je charakteristická přítomností několikanásobného rématu (viz výše), jednak tím, že jednotlivé části rozštěpeného rématu zakládají samostatné tematické posloupnosti, a jsou tedy formálně začleněny do struktury V1; „rozvíjení posloupnosti štěpeného rématu se liší od jiných posloupností tím, že je to vlastně posloupnost vyššího stupně, či přesněji posloupnost rámcová, která uspořádává dvě nebo více posloupností (nižšího stupně)“ (Daneš 1968: 135), je to tedy posloupnost strukturně složitější než dosavadní představené varianty. Zároveň ale není podmínkou, aby byly v širším kontextu rozvíjeny všechny části rozštěpeného rématu, ani není omezen typ tematické posloupnosti, který musí jednotlivé dílčí posloupnosti naplňovat, dokonce se schémata jednotlivých dílčích posloupností mohou navzájem lišit. Stejně tak neexistují omezení týkající se rozsahu prostoru, který má být jednotlivým dílčím posloupnostem věnován.

Jedním z příkladů, které Daneš (1968) k tomuto schématu uvádí, je kontext *Na počátku 17. století položili základ novému rozvoji astronomie dva velcí mužové. Jan Kepler založil teoretickou astronomii. Ukázal, jak je možno z pozorování odvodit... Galileo Galilei založil mechaniku. Svými pokusy...* Tematická část ve větě je velice široce vymezena, její podmínkou je jakákoli vazba mezi tématy a rématy (explicitní nebo vydedukovatelná na základě kontextu) následujících výpovědí. Není tedy podmíněna koreferenčním vztahem mezi entitami v tom smyslu, v jakém jsme koreferenci vymezili.

V případě tohoto typu tematické posloupnosti je více než v jakémkoli jiném schématu důraz kladen na signalizaci toho, že réma je dvoučlennou strukturou (k tomu

se užívá zejména dvojčlenných prostředků typu *jednak – jednak, na jedné straně – na druhé straně* apod.) a signalizaci toho, ke které tematické posloupnosti daná výpověď patří (k tomu slouží zejména opakování plnovýznamového vyjádření nebo výrazy typu *pokud jde o, v prvním případě – v druhém případě* apod.).

### 3.2.4 Tematické odvozování

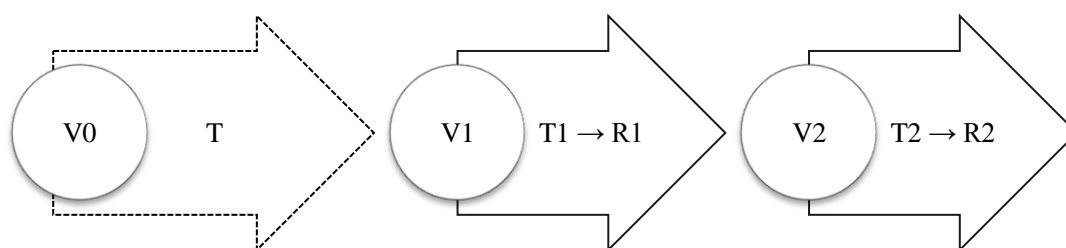


Schéma G: tematické odvozování

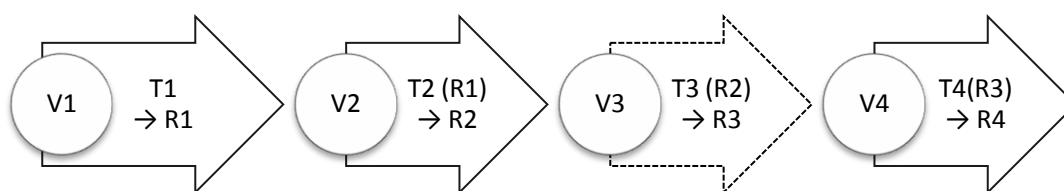
Základním prvkem výstavby schématu charakterizovaného jako tematické odvozování je přítomnost hypertématu, z něhož jsou jednotlivá dílčí témata odvozována. Zcela zřetelně je jasné, a i Daneš si to uvědomuje, že je velice úzká vazba mezi průběžným tématem a tematickým odvozováním. Rozdíl ozřejmuje Daneš takto: „Z naší charakteristiky posloupnosti s tematickým odvozováním vyplývá, že u tohoto typu má každá výpověď své vlastní téma (byť odvozené z tématu hlavního, které popř. může být ve výpovědi také vyjádřeno nebo na ně odkázáno). [...] Naproti tomu u typu s průběžným tématem, se ukazuje, že toto průběžné téma (které je jistou analogií nadřazeného základního tématu u typu tematického odvozování) je tématem všech výpovědí a že tyto výpovědi žádné další své téma nemají; vše, co tyto odpovědi obsahují navíc kromě tohoto průběžného tématu, patří už k rématu.“ (Daneš 1968: 137) Příkladem, který by měl osvětlit rozdíly mezi tímto typem tematické posloupnosti a průběžným tématem, je kontext: *Záškrť (diphtheria) je infekční onemocnění. Působí je corynebacterium diphtheriae. Šíří se kapénkovou infekcí přímým stykem s nemocným, častěji však bacilonosičem, nebo nepřímou předměty potřísněnými hlenem. Inkubační doba je 2 až 5 dnů. Pořadí uváděných informací není závazné, nicméně „volba odvozených témat [je] dána stabilními prvky základního*



tématu a jejich sled může být – v daném oboru anebo alespoň v dané práci či stati – víceméně ustálen“ (ibid: 138), velice často bývá ustálena forma zejména v případě chronologického postupu. Pro srovnání uvedeme na tomto místě i kontext, který Daneš označuje za průběžné téma: *V oboru izolátorů se věnuje velká pozornost tzv. ferroelektrikům. Tyto látky mají schopnost měnit elektrickou energii v mechanickou a naopak, čehož lze využít k výrobě jednoduchých, trvanlivých a účinných přenosek, reproduktorů apod. Užívá se jich také pro tzv. paměť automatických počítačů. Pomocí těchto látek lze také až několikrát zmenšit rozměry elektrických kondenzátorů.*

Je tedy patrné, že vymezení tématu je v případě průběžného tématu specifitější, kdežto v tematické posloupnosti tematického odvozování je vymezeno široce, tak, že z něj je možné odvozovat jednotlivá dílčí témata.

### 3.2.5 Tematický skok



**Schéma H: tematický skok**

Poslední typ tematické posloupnosti je považován za zvláštní a jeho specifitnost spočívá v tom, že dojde v rámci schématu klasické tematizace rématu k vynechání jednoho stupně, který „se jeví jako samozřejmý (a snadno se doplní ze souvislosti)“ (Daneš 1968: 138). Konkrétní explicitní kontext u tohoto typu tematické posloupnosti Daneš (1968) neuvádí, nicméně odkazuje na texty typu výkladu o výrobě hedvábí, v nichž není explicitně uvedeno předení jako jedna z fází této výroby, nicméně se „přádelny hedvábí (ve výpovědi *Přádelny hedvábí jsou hlavně v městech A, B a C.*) objeví nikoli nejprve jako réma, nýbrž rovnou jako téma, skokem (ihned po výkladu o tom, jak se získává surovina k výrobě hedvábí)“ (ibid: 138).

Pět typů tematických posloupností představuje ideální případy, a tedy nutně abstraktní konstrukty. Daneš je zakládá na modelových neproblematických příkladech. Mezi jednotlivými tematickými posloupnostmi není zřetelný předěl v tom smyslu, že by jeden text byl vystavěn pouze z jednoho typu tematické posloupnosti. Naopak, takový text by byl velice nepřirozený a příznakový. Neutrální výstavba textu je charakteristická tím, že se jednotlivé typy tematických posloupností kombinují a začleňují do sebe. Takovýmto způsobem mohou být tematické posloupnosti komplexní a rekurzivní: „výpověď [...] je vlastně komplexní, [když] obsahuje v sobě zkráceně (kondenzovaně) výpovědi dvě“ (ibid: 131). Příkladem takovéto výpovědi může být *Surové železo zbavené síry postupuje dále nepřetržitým tokem do otáčivé válcové pece, kam se přivádí kyslík a dodává vápno.*, která v sobě zahrnuje výpovědi 1) *Surové železo zbavené síry postupuje dále nepřetržitým tokem do otáčivé válcové pece.*, 2) *Do této otáčivé válcové pece se přivádí kyslík a dodává vápno.* Tato Danešova interpretace by mohla být rekurzivně rozšířena i dále, neboť i v rámci první výpovědi nacházíme kondenzaci *surové železo zbavené síry*, kterou bychom mohli analyzovat jako 1) *Surové železo je zbaveno síry.*, 2) *Toto/Surové železo postupuje dále...* V takovéto Danešově analýze spatřuji nekonzistentní přístup k souvětným výpovědím, neboť není zřejmé, jak vymezit, co je kondenzací a co jí není, tedy která část výpovědi je a která už není považována za komplexní části výpovědi.

Na konci svého článku z roku 1968 podává Daneš stručný přehled způsobů, jakými lze vyjadřovat témata ve vztahu k předchozímu rématu. Vzhledem k tomu, že formy vyjadřování témat nejsou předmětem této práce, uvádím jen přehled možností: a) téma může být vyjádřeno parafrází předchozího rématu, b) téma je zřetelně obsaženo v předchozím rématu, c) téma je implikováno v předchozím rématu méně zřetelně, vzniká zdání tematického skoku, d) téma kontrastní. Kromě toho mohou do tematické posloupnosti zasahovat různé odbočky a vsuvky.

### 3.3 *Ludmila Zimová*

Druhým přístupem, který reprezentuje větněčlenskou perspektivu uplatňovanou v této práci, je přístup Ludmily Zimové (1953–). Ve své práci *Způsoby vyjadřování větných členů v textu: konkurence pojmenování, pronominalizace a elize* (1994) se věnuje podmětu, předmětu a příslovečnému určení. Jak je z názvu patrné, zabývá se především srovnáním různých způsobů vyjádření větných členů v textu. Takovýto přístup poskytuje možnost reflektovat i případné významové rozdíly, které z různé realizace daného větného členu vyplývají.

Co se týče užívaných označení, Zimová pracuje s dvoučlennými posloupnostmi, rozlišuje antecedentní vs. postcedentní vyjádření, tj. první a druhou větu, v nichž jsou koreferenční vyjádření (antecedent a postcedent).<sup>20</sup> Z celé práce se zaměříme na zpracování podmětu a předmětu, neboť ty jsou pro tuto práci klíčové.

#### 3.3.1 *Podmět*

Kapitoly věnované podmětu jsou nejrozsáhlejší. Za primární variantu realizace podmětu v textu považuje Zimová především elizi, tj. nevyjádřený podmět. Podmínky, které musí být splněny, aby mohla nastat elize podmětu, jsou zřejmé: „subjekt nebývá v češtině vyjádřen, je-li znám ze situace, v níž jazykový projev probíhá, nebo z předchozího jazykového kontextu, anebo ukazuje-li na něj svým mluvnicky shodným tvarem predikát (finitní slovesný tvar)“ (Zimová 1994: 59). Pokud tyto podmínky splněny nejsou, je v textu podmět povrchově přítomen ve formě plného vyjádření nebo ve formě zájmena.

Zimová přijímá Greplův–Karlíkův (1986: 239) předpoklad, že „užití osobního zájmena ve funkci podmětu je vždy nějak motivováno“ (Zimová 1994: 59), tedy že vyjadřování podmětu v textu není primární a neutrální variantou, ale variantou příznakovou. Zároveň „[...] činitelé motivující vyjádření podmětu osobním zájmenem jsou různé povahy. [...] Za základní lze považovat faktory [...]: 1. jestliže osoba není dostatečně určena koncovkou přísudkového slovesa, 2. má-li být subjekt rozvit, 3. je-li subjekt zdůrazněn, 4. při vytčení do protikladu.“ (ibid: 60)

<sup>20</sup> Vzhledem k velkému množství chyb v originálním textu jsme se rozhodli v citacích chybné pasáže (chybné koncovky, překlapy apod.) opravovat a uvádět gramaticky správné varianty. Jiné změny prováděny nebyly.

Zimová zavádí terminologické rozlišení realizace podmětu. Podmět realizovaný osobními zájmeny první a druhé osoby nazývá v souladu s dosavadní odbornou literaturou zájmenný podmět, tedy podmět vyjádřený zájmeny, zatímco podmět realizovaný zájmenem třetí osoby, ukazovacími zájmeny *ten* a *to* označuje jako podmět pronominalizovaný, tedy jako „výsledek kontextově podmíněného způsobu realizace subjektové pozice – pronominalizace“ (ibid: 60). Tato diplomová práce se zaměřuje na zájmeno třetí osoby a jeho roli. Termínu pronominalizace užíváme ve stejném smyslu jako Zimová (1994) ovšem s tím rozdílem, že ho užíváme pouze pro zájmena osobní a přivlastňovací. Ne z toho důvodu, že pro ukazovací zájmena ho užívat nechceme, ale proto, že ukazovací zájmena nejsou předmětem našeho zkoumání (viz oddíl Zájmenná koreference, 4.2.5).

Co se týče pronominalizace, je podle Zimové motivována čtyřmi faktory: „snahou autora jazykového projevu zdůraznit již zmíněný objekt řeči, důvody rytmickými, emocionálním postojem autora jazykového projevu ke sdělované skutečnosti, autorovým individuálním územ“ (Zimová 1994: 77). Tyto faktory se projevují především v konkurenci vyjádření podmětu osobním zájmenem *on* a *elize*.

V některých případech je povrchová realizace podmětu plným pojmenováním obligatorní. Povrchová realizace plného pojmenování může sloužit jako prostředek eliminace nejednoznačnosti v potenciálně nejednoznačných kontextech (např. *U mlýna před vraty stála socha svatého Jana Nepomuckého mezi dvěma lipami: tam sedávala v neděli po obědě panímáma, někdy také žernovská kmotra, co zůstávala na stráni, a Mančinka. Jakmile zahlédly babičku s dětmi okolo strouhy přicházet, běžela jim Mančinka vstříc.*). V případech, kdy dochází ke změně podmětu oproti předchozí větě, slouží plné pojmenování jako jasný signál této změny a eliminuje nepochopení nebo případné dezinterpretace. V případech, kdy dochází k opakovanému výskytu entity v textu po delší době, slouží plné pojmenování jako prostředek jednoznačného ztotožnění entity, jejíž výskyt je textově vzdálený, s entitou, která do textu právě znovu vstupuje.

### 3.3.2 *Předmět*

Druhým analyzovaným větným členem je předmět. Co se týče předmětu, není prostor jemu v textu věnovaný tak rozsáhlý, i přesto, že se jedná o druhý nejfrekventovanější větný člen. V úvodní části věnované předmětu Zimová vymezuje vztah mezi podmětem a předmětem

z hlediska prostředků vyjadřování tohoto větného členu v textu: „Možnost elize objektu [...] se nám jeví ve srovnání s elizí subjektu výrazně menší. To je dáno především tím, že neobsazení subjektové pozice je zpravidla kompenzováno kongruenčními koncovkami verba finita v predikátu. Realizace subjektu zájmenem, resp. nutnost vyjádřit subjekt plnovýznamovým vyjádřením je vyvolána [...] působením určitých specifických podmínek. U objektové [...] pozice se naopak jeví jejich realizace – ať už plnovýznamovým pojmenováním nebo zájmeným výrazem – jako základní.“ (Zimová 1994: 92)

Elizi předmětu Zimová tedy považuje za netypickou formu vyjádření předmětu, vylučuje ji v případech, kdy se zabývá konkurencí elize a pronominalizace ukazovacím zájmenem *ten*, a to zejména proto, že „postcedent vyjádřený ukazovacím zájmenem je zdůrazněn, je sdělně závažný, a proto nemůže být elidován“ (ibid: 92).

Zastoupení předmětu ve větě ve formě elize podléhá podle Zimové několika vlivům: počtu objektových valenčních pozic, typu předmětu (přímý vs. nepřímý) – např. „[...] u vícevalenčních predikátů může být elize jednoho z objektů závislá na elizi objektu druhého.<sup>21</sup>

Nevyjádření některých objektů chápe Zimová ne jako elizi objektu, ale jako užití odlišného predikátu. Srovnajme vyjádření *Puk směřuje do levého horního rohu naší branky, ale Holeček o/ho pohotově chytá a padá na led.* a *Puk směřuje do levého horního rohu naší branky, ale Holeček chytá pohotově, a tak se skóre nemění.* (ibid: 121). V prvním příkladě jde podle Zimové o elizi objektu tak, jak jsme o ní uvažovali doposud. V příkladě druhém naopak nejde o elizi objektu, neboť se v této větné struktuře uplatňuje jiný predikát, „*chytat* ve významu ‘být v brance, být brankářem’“ (ibid: 122), které vůbec nemůžeme považovat za objektové.

<sup>21</sup> Zdá se, že tzv. nepřímý předmět v dativu podléhá elizi snáze. Dojde-li k elizi akuzativního předmětu přímého, eliduje se současně i předmět nepřímý v dativu. Elize tohoto předmětu však elizi akuzativního předmětu nezpůsobí.“ (Zimová 1994: 116) S tímto tvrzením nemůžeme než souhlasit, Zimová uvádí přesvědčivé příklady: *Slíbil, že mi knihu vrátí v pátek, ale nevrátil 0[mi] ji/0.* vs. *Slíbil, že mi knihu vrátí v pátek, ale nevrátil mi \*0[ji].* Zároveň jsou jakékoli elize objektu vyloučeny v případech typu *Vysoukám se ze svého spacího pytle, vstojí na žebříku ho/\*0 sroluju, přetáhnu přes něj plachtu a vydám se na cestu.*, kdy „mezi antecedentem a postcedentem není úplný soulad co do pádu a větného členství [...], i když jsou věty spojeny vztahem koordinace.“ (ibid: 117)

Další možností vyjádření předmětu je povrchová realizace plného pojmenování. K této variantě ale Zimová ve srovnání s elizí v odpovědích na otázky (např. *Zavezl jsi auto do garáže?*) poznamenává, že „Elizi [...] lze v těchto případech považovat za jev zcela běžný, někdy stylisticky příznakový. [*Zavezl.*] Plná realizace větné struktury [typu *Auto jsem do garáže zavezl.*] by byla [...] adresátem pocíťována jako cosi neobvyklého.“ (ibid: 115).

Pronominalizace předmětu, jak je patrné již z několika uvedených příkladů, konkuruje elizi: *Poirot se nápadně rychle vztyčil z pokleku a několika drobnými kroky se octl u dveří. Otevřel je/0, okamžik stál nehnutě, pátravě hleděl ven a pak je/0 zase zavřel.* Kromě toho ale o pronominalizaci můžeme spolu se Zimovou uvažovat i v konkurenci vyjádření demonstrativem *ten* a osobním zájmenem, a to v případech definovaných na základě morfologických vlastností antecedentu a postcedentu: „Pro pronominalizaci předmětu ukazovacím zájmenem *ten* nebo *on* mají rozhodující význam tyto mluvnické kategorie: u pojmenování v antecedentu kategorie mluvnického rodu a čísla, u zájmenného postcedentu navíc ještě kategorie pádu.“ (ibid: 93).

Vyjádření demonstrativní bývá užíváno v případech, kdy a) postcedent plní funkci zdůrazňovací a je vytčen na začátek věty, b) postcedent stojí v kontaktním postavení, neboť demonstrativum *ten* „odkazuje k pojmenování bezprostředně předcházejícímu, které bývá zpravidla rématem antecedentní věty“ (ibid: 93), c) postcedent stojí v distančním postavení, ale jeho interpretace je jednoznačná. Je ale zřejmé, že existuje velké množství situací, kdy si vyjádření demonstrativem a osobním zájmenem konkurují. To jsou případy, kdy o volbě demonstrativního nebo osobního zájmena rozhodují postavení v důrazové pozici, vícenásobnost antecedentu a rytmičné aspekty. V postavení důrazovém se podle Zimové uplatňuje zájmeno ukazovací, zatímco zájmeno osobní je typické pro postavení nedůrazové. Stejně tak existuje preference k užití demonstrativa v případech, kdy u vícenásobného antecedentu mluvčí odkazuje k poslednímu členu. Rytmičné aspekty se stávají klíčovými zejména ve srovnání pádů prostých a předložkových. „Přítomnost slabičné předložky – a tedy vytvoření minimálně dvouslabičného taktu s první slabikou přízvučnou – znamená výrazný posun ve vztahu konkurence obou zájmen ve prospěch zájmena osobního. Je to zřejmé především v těch pádech, v nichž zvuková nevýraznost a

krátkost tvarů osobního zájmena vede k preferenci zájmena ukazovacího.<sup>22</sup> *Ty pospícháš? Snad ne k ženě? Neboj se. O tu/ni se ti pan správce postará. // [...] Tu/ji doma stejně nenajdeš.*“ (ibid: 97). Z důvodů rytmických jsou tedy tvary demonstrativa preferovány v pádech nepředložkových, zatímco tvary osobního zájmena v pádech předložkových.

### 3.3.3 Konkurence vyjadřovacích prostředků

Každý větný člen v textu má tendenci být primárně vyjadřován jednou formou realizace (elize, pronominalizace, plná fráze). Zimová se také věnuje preferencím navazování na jednotlivé větné členy v závislosti na jejich formě realizace v textu. Nyní představíme to, jaké způsoby realizace v textu jsou pro jednotlivé větné členy základní.

Zdá se, že základním kritériem, které odlišuje podmět od předmětu, je možnost elize. Elize je typickou formou realizace podmětu, zatímco v případě předmětu je naopak formou spíše okrajovou. Pronominalizace je základní formou vyjadřování větných členů v textu v případě, kdy nenastává referenční nejednoznačnost a zároveň autor nechce především ze stylových důvodů opakovat plnovýznamové pojmenování. Plnovýznamové pojmenování se v postcedentní pozici v textu může realizovat třemi způsoby: „opakování téhož pojmenování, vyjádření pojmenováním synonymním a vyjádření pojmenování s jiným významem“ (Zimová 1994: 131). Tyto tři varianty nepovažuje Zimová za dostačující kritéria pro postizení všech možností realizace větných členů v textu, neboť shledává „rozdíl mezi genericky a individualizovaně užitým pojmenováním“ (ibid), který považuje za důkaz toho, že ani opakování téhož pojmenování nemusí odkazovat ke stejnému objektu řeči.

Za konkurenční vyjádření Zimová považuje v případě podmětu zájmenné realizace a elizi, v případě předmětu zájmennou realizaci *on* a *ten*: „Ke konkurenci pronominalizace *ten/on* a elize dochází jen v pozici subjektové, v pozici objektu [...] konkurují toliko zájmena *ten* a *on*, a to pouze v postavení důrazovém. V nedůrazovém postavení se uplatňuje toliko zájmeno osobní.“ (Zimová 1994: 164) Užití plnovýznamového pojmenování v postcedentu je velice omezené, dochází k němu v podstatě jen v případech referenční nejednoznačnosti při užití jiných forem vyjádření: „Je-li v antecedentní větě

<sup>22</sup> Podle našeho názoru může důležitější než zvuková nevýraznost být spíše důrazové postavení. Ale bylo by potřeba provést podrobnější zkoumání, abychom mohli činit takové závěry.

více substantivních pojmenování téhož gramatického rodu a čísla, která by při pronominalizace, resp. elizi [...] v postcedentní větě přicházela v úvahu jako antecedent a která by i z hlediska lexikálněsémantické spojitelnosti slov mohla plnit funkci subjektu postcedentní věty, užije se k realizaci postcedent plnovýznamového pojmenování.“ (ibid: 165)

Můžeme tedy shrnout rozdíly mezi podmětem a předmětem, které Zimová uvádí a které jsme zmiňovali výše, do přehledné, ale zjednodušené tabulky B:

	podmět	předmět
<i>základní forma realizace</i>	elize	pronominalizace nebo plnovýznamové pojmenování
<i>elize</i>	ano	ne
<i>pronominalizace</i>	ano	ano
<i>zájmeno osobní on</i>	3. osoba	nedůrazové postavení předložkové pády
<i>zájmeno ukazovací ten</i>	ano	důrazové postavení nepředložkové pády
<i>plnovýznamové pojmenování</i>	ne	ne
<i>identifikátor ten</i>	ne	ano (identifikace) ne (jiné prostředky)

**Tabulka B: podmět, předmět**

### 3.3.4 Topik – fokus

Ludmila Zimová se zabývá především větněčlenskou platností, tj. aktuálněčlenská distinkce není v textu explicitně tematizována. Informace o postavení v topikou nebo fokusu můžeme nalézt většinou jen v případech srovnávání konkurenčních způsobů realizace větných členů v textu (pronominalizace, elize, opakování plné jmenné fráze), kde existují rozdíly podle toho, zda je člen v postavení topikovém, nebo fokusovém. I přesto ale nacházíme některé postřehy týkající se aktuálního členění zejména ve vztahu k podmětu a předmětu.

Podle Zimové se „aktuální členění jako činitel ovlivňující vyjádření podmětu osobními zájmeny 1. a 2. osoby projevuje v citově neutrálních výpovědích především tehdy, je-li podmět rématem výpovědi“ (Zimová 1994: 60). Postavení podmětu v rématu tedy není primární realizací, je prostředkem signalizujícím odlišnou pragmatickou funkci výpovědi, než je prosté neutrální vyjádření. Srovnajme výpovědi typu *Když dostanu jablko*



*0, budeš nejslavnějším vojevůdcem. a Když dostanu jablko já, budeš nejslavnějším vojevůdcem.*, které Zimová uvádí. Ve výpovědi prvního typu, kde není podmět explicitně povrchově vyjádřen, je pro srovnání zaznamenáno, kde je jeho pozice v rématu. Ale je těžko představitelné, že by první uvedenou větu někdo interpretoval jako výpověď s rematickým podmětem. Primární variantou je podmět realizovaný (i nevyjádřený) v pozici tématu (*Když já/0 dostanu jablko, budeš nejslavnějším vojevůdcem.*) a to by pravděpodobně byla preferovaná inferovaná varianta v případě tzv. nevyjádřeného podmětu.

V případě utváření koreferenčních řetězců pronominalizací Zimová tvrdí, že „působí [...] tendence, aby byl podmět, dochází-li ve větné posloupnosti k jeho změně, zřetelně signalizován (nestačí-li k tomu koncovky predikátu, je nutno vyjádřit jej slovně)“ (ibid: 71). Pokud je tedy podmět v průběhu textu z hlediska reference změněn, je nutné tuto změnu explicitně signalizovat. Kromě toho si Zimová všímá dalšího faktu: „Osobní zájmeno ve funkci subjektu [...] napomáhá vyjádřit to, že antecedentem není pojmenování v subjektu věty předchozí, nýbrž její jiný substantivní člen. Réma antecedentní věty [...] je tématem věty následující a je pronominalizováno. Elize subjektu v postcedentní větě by s vysokou mírou pravděpodobnosti odkazovala k subjektu věty antecedentní.“ (ibid: 75n) Z toho jasně vyplývá, že podle Zimové je preferovanou variantou utváření koreferenčního řetězce řetězec s antecedentem v podmětovém postavení.

K utváření tematických posloupností v textu se Zimová vyjadřuje takto: „Antecedent bývá zpravidla jádrem výpovědi, jeho postcedent východiskem výpovědi následující.“ (ibid: 81) Z toho vyplývá, že za základní metodu výstavby textu považuje Zimová v souladu s Danešem (1968) tematizaci rématu.

V závěru věnuje Zimová jeden odstavec postavení antecedentu zájmena v topiku a fokusu v případě vyjádření anaforu koreferenčního řetězce plnovýznamovým pojmenováním bez identifikátoru *ten* ve srovnání s demonstrativním *ten*: „Lze konstatovat, že se projevuje silná tendence, aby postcedent realizovaný obecným pojmenováním bez identifikátoru zaujímal pozici tematickou (v objektivním pořádku slov), ale také aby téma bylo zřetelně signalizováno adjektivním ukazovacím zájmenem *ten*. Užití identifikátoru *ten* v postcedentu rematickém se jeví závazné jen tehdy, je-li mezi antecedentem a postcedentem velká vzdálenost.“ (ibid: 168)

Můžeme tedy shrnout, že podle Zimové je nejčastější (preferovanou) variantou utváření koreferenčních řetězců a) navazování na rematickou část antecedentní věty, b) navazování na podmět antecedentní věty.

### 3.4 Centering

Centering je teorie vytvořená americkými lingvisty na přelomu 70. a 80. let, programová stat' Grosz – Joshi – Weinstein: *Towards a computational theory of discourse interpretation* byla napsána, ale nepublikována v roce 1986, později bylo vydáno velké množství navazujících prací (Walker 1994, Grosz – Weinstein – Joshi 1995, Walker – Joshi – Prince<sup>23</sup> 1998 atd.). Za primární zdroj je považován právě text WJP (1998), protože podává základní přehled o základech a metodologii teorie centeringu. Kromě toho je klíčovým textem i Brennan – Friedman – Pollard<sup>24</sup> (1987) a Sidner (1981), kteří se zaměřují na uchopení a fungování zájmen z hlediska centeringu. Teorie se zaměřuje na modelování lokální koreference<sup>25</sup> a vychází z předpokladu, že různé sémantické části výpovědi jsou různě důležité z hlediska informační přínosnosti. V praxi je teorie centeringu využívána např. v části HPSG věnované zpracování pragmatiky, konkrétně jde o algoritmus vázání zájmen (*pronoun binding*), viz dále.

#### 3.4.1 Centra pozornosti

Teorie centeringu chápe text jako ucelenou jednotku, která je dále členěna na soubor tzv. segmentů, které v podstatě odpovídají odstavcům (označovaným D), segmenty poté na jednotlivé výpovědi (označované U). „Centeringový model je velice jednoduchý. Diskursy se skládají z konstitutivních segmentů, z nichž každý je reprezentován jako část modelu diskursu. Centra jsou sémantické entity, které jsou součástí modelu diskursu pro každou výpověď v segmentu.“<sup>26</sup> V rámci každé výpovědi jsou vyčleněny sémantické entity, tedy centra pozornosti (označované C), které náleží do tří typů.

<sup>23</sup> Dále jen WJP.

<sup>24</sup> Dále jen BFP.

<sup>25</sup> „centering is defined on the basis of local feature sof utterances“ (WJP 1998: 22)

<sup>26</sup> „The centering model is very simple. Discourses consists of constituent segments and each segment is represented as part of a diskurse model. Centers are semantic entities that are part of the discourse model for each utterance in a discourse segment.“ (WJP 1998: 2)

1. Cf (*forward-looking center*) – centra pozornosti, která v textu „hledí vpřed“ a poskytují možnost založit koreferenční řetězec (z opačné perspektivy jsou to potenciální antecedenty, které anafor hledá v předchozích částech textu): „Soubor Cf reprezentuje entity diskursu, které jsou do něj uvedeny prostřednictvím výpovědi  $U_i$  v segmentu diskursu D.“,<sup>27</sup> BFP uvádějí: „S každou výpovědí je asociován soubor Cf, který zahrnuje entity diskursu, které jsou přímo realizovány nebo realizovány<sup>28</sup> jazykovými výrazy.“<sup>29</sup>
2. Cp (*preferred center*) – v rámci center pozornosti Cf je Cp v hierarchii nejvýše postavené; kritériem pro určení této hierarchie je a) postavení ve větnečlenské funkci (Cp se stává nejpravděpodobněji Po, poté Pt a poté další větné členy)<sup>30</sup>, b) slovosled v případě jazyků, kde je slovosled klíčový pro přiřazení gramatických rolí mimojazykovým entitám: „Soubor Cf je hierarchizován podle salience. Toto uspořádání je relativní. Nejvýše postavený člen souboru Cf je označován jako PREFEROVANÉ CENTRUM, Cp. Předpokládáme, že toto PREFEROVANÉ CENTRUM se stane Cb následující výpovědi. V některých případech je Cp vyjádřením toho, čeho se týkala předchozí výpověď diskursu, jejím Cb, ale neplatí to vždy.“<sup>31</sup> „Řadíme položky v souboru Cf na základě obligatornosti gramatického vztahu podřízených funkcí predikátu.“<sup>32</sup>

<sup>27</sup> „The set of FORWARD-LOOKING CENTERS, Cf( $U_i$ , D), represents discourse entities evoked by an utterance  $U_i$  in a discourse segment D.“ (WJP 1998: 3)

<sup>28</sup> K distinkci realizace a přímé realizace viz dále.

<sup>29</sup> „With each utterance  $U_n$  is associated a list of *forward-looking centers*, Cf( $U_n$ ), consisting of those discourse entities that are *directly realized* or *realized* by linguistic expressions in the utterance.“ (BFP 1987: 155)

<sup>30</sup> V případě češtiny je centrem pozornosti Cp prakticky ve všech případech podmět.

<sup>31</sup> „The set of FORWARD-LOOKING CENTERS, Cf, is ranked according to discourse salience. This rating is a partial order. The highest ranked member of the set of forward looking center is referred to as the PREFERRED CENTER, Cp. The PREFERRED CENTER represents a prediction about the Cb of the following utterance. Sometimes the Cp will be what the previous utterance of discourse was about, the Cb, but this is not necessarily the case.“ (WJP 1998: 3)

<sup>32</sup> „We rank the items in Cf by obliqueness of grammatical relation of the subcategorized functions of the main verb.“ (BFP 1987: 156)

3. Cb (*backward-looking center*) – centrum pozornosti Cb odkazuje k předchozí části textu a je v rámci zkoumání koreference považováno za anafor; centrem pozornosti Cb se stává nejvýše stojící element předchozí výpovědi; zároveň má Cb několik omezení: mohli bychom ho přirovnat k topiků s tím omezením, že jako Cb může být ve výpovědi označena pouze jedna entita, ať už explicitně zmíněná, nebo ne: „Cb je speciálním členem Cf, který reprezentuje entitu diskursu, které se výpověď U<sub>i</sub> nejvíce věnuje. Mohli bychom ho přirovnat k tomu, co jiní (Reinhart, 1981; Horn, 1986) označují jako topik: Cb propojuje aktuální výpověď s předchozím kontextem.“<sup>33</sup>

Tato centra pozornosti se typicky vyskytují v každé výpovědi,<sup>34</sup> jsou vzájemně provázána a jejich vztahy jsou specifikovány několika následujícími zásadami týkajícími se vymezení Cf, Cp a Cb v rámci výpovědi (1., 2.) a pravidly upravujícími vztahy mezi Cf, Cp a Cb sousedních výpovědí (3.–5.). Zásady 1.–3. jsou označovány jako omezení (*constraints*), zásady 4. a 5. jako pravidla (*rules*).<sup>35</sup>

1. Každý prvek souboru Cf musí být ve výpovědi realizován.
2. Ve výpovědi je maximálně jedno centrum pozornosti Cb. – Toto tvrzení předpokládá, že existuje jedna entita výpovědi, která je centrální, a tou je Cb. WJP uvádějí, že psycholingvistické výzkumy tento předpoklad potvrzují. V rámci empirického výzkumu bylo doloženo, že v jedné výpovědi neexistuje více Cb než jedno, ale že výpověď nemusí nutně Cb mít, tj. může sestávat jen z Cf a Cp.
3. Centrem pozornosti Cb v aktuální výpovědi se stává to centrum pozornosti Cf, které je nejvýše postaveným elementem Cf předcházející výpovědi, tj. jejím Cp, realizovaným v aktuální výpovědi. – Realizací není myšleno pouze

<sup>33</sup> „The BACKWARD-LOOKING CENTER, Cb (U<sub>i</sub>, D) is a special member of the Cf, which represents the discourse entity that the utterance U<sub>i</sub> most centrally concerns, similar to what is elsewhere called the ‘topic’ (Reinhart, 1981; Horn, 1986): The Cb entity links the current utterance to the previous discourse.“ (WJP 1998: 3)

<sup>34</sup> Je zřejmé, že existují variace počtu Cf v jednotlivých výpovědích.

<sup>35</sup> Volně parafrázováno podle WJP 1998: 3.

povrchové vyjádření (např. zájmenem), ale také povrchové nevyjádření, přítomnost v hloubkové struktuře (např. nevyjádřený podmět). Specifickým případem realizace je tzv. přímá realizace (*relation direct realize*), ke které dochází v případě, že Cb zastupuje celou předchozí výpověď (např. v kondenzované podobě). Je tedy zřejmé, že Cb(U<sub>i</sub>) a Cp(U<sub>i-1</sub>) jsou dva vzájemně provázané elementy.

4. Cf první výpovědi se v druhé výpovědi stává Cb, pokud je pronominalizované. – V této zásadě je reflektován fakt, že Cb bývá často (i v češtině) pronominalizováno nebo elidováno. V případě, že ve druhé výpovědi je více pronominalizovaných entit, je realizací Cb jedna z nich. Toto pravidlo bývá nazýváno zájmené pravidlo.
5. Existuje hierarchie typů přechodů mezi výpověďmi založená na míře koherence textu: 1. *continue*, 2. *retain*, 3. *smooth-shift*, 4. *rough-shift*. – Toto pravidlo předpokládá, že některé typy přechodu zajišťují vyšší míru koherence (*continue*) než jiné (*rough-shift*).

### 3.4.2 Typy přechodů

Výstavba textu z hlediska aktuálního členění je v teorii centeringu zpracována pomocí tzv. přechodů čtyř různých typů, *continue*, *retain*, *smooth-* a *rough-shift*.<sup>36</sup>

Následující tabulka C zachycuje typologii těchto přechodů na základě dvou kritérií vycházejících ze vzájemných vztahů mezi centry pozornosti uvedených tří typů. Prvním kritériem, které vymezuje přechody *continue* a *retain* vůči přechodům *smooth-shift* a *rough-shift*, je vztah Cb ve výpovědi aktuální (U<sub>i</sub>) a jí bezprostředně předcházející (U<sub>i-1</sub>).<sup>37</sup> V případě, že Cb aktuální výpovědi je shodné s Cb výpovědi předcházející, mluvíme o přechodech *continue* a *retain*, které bychom tedy mohli považovat za reprezentanty Danešovy tematické posloupnosti typu průběžného tématu. V případě, že se Cb promění, mluvíme o přechodech *smooth-* a *rough-shift*. Tyto přechody bychom

<sup>36</sup> S označeními typů přechodů v teorii centeringu pracuji jako s termíny a používám anglická označení. České překlady (pokračování, ponechání, hladký a hrubý posun) jsou podle našeho názoru zavádějící a nejsou ustálené.

<sup>37</sup> V případě první výpovědi v diskursu je určení Cb obtížné, neboť neexistuje předchozí výpověď, s níž by bylo možno srovnat Cb. V takovém případě platí Cb(U<sub>i-1</sub>) = ?.

mohli spojit s Danešovou tematickou posloupností typu tematizace rématu. Druhým kritériem je vztah  $C_b$  a  $C_p$  v aktuální výpovědi. Pokud platí vztah  $C_b = C_p$ , jedná se o přechody *continue* a *smooth-shift*, pokud tento vztah neplatí, jedná se o přechody *retain* a *rough-shift*. Zjednodušeně řečeno bychom toto kritérium mohli parafrázovat jako postavení podmětu v topiku.<sup>38</sup> V tom případě by typy, kde  $P_o$  stojí v  $T$ , tedy *continue* a *smooth-shift*, reprezentovaly neutrální nastavení kontextu. Typy, kdy dochází k jinému postavení  $P_o$  než topikovému, tj. *retain* a *rough-shift*, by zastupovaly uspořádání příznakové. V tabulce C je zachycena kombinace těchto dvou kritérií.

	$C_b(U_i) = C_b(U_{i-1})$ $C_b(U_{i-1}) = [?]$ <sup>39</sup>	$C_b(U_i) \neq C_b(U_{i-1})$
$C_b(U_i) = C_p(U_i)$	<i>continue</i>	<i>smooth-shift</i>
$C_b(U_i) \neq C_p(U_i)$	<i>retain</i>	<i>rough-shift</i>

**Tabulka C: „tematické“ posloupnosti v teorii centeringu**

Typ *continue* se vyznačuje zachováním  $C_b$  předchozí výpovědi, které je v aktuální výpovědi  $C_p$ . Z toho důvodu je koherence textu v případě navazování typu *continue* nejvyšší. Typ *retain* také zachovává  $C_b$  z předchozí výpovědi, nicméně už není v pozici  $C_p$ . Z toho důvodu je koherence textu v případě tohoto navazování nižší. Přechody označované jako *shift* se vyznačují tím, že  $C_b$  z předchozí věty nezachovávají, ale  $C_b$  mají, v případě přechodu *smooth-shift* je jím  $C_p$  aktuální výpovědi, v případě přechodu *rough-shift* je jím jiné centrum pozornosti než  $C_p$ . „Distinkce mezi typy přechodů *smooth-shift* a *rough-shift* byla poprvé zavedena v BFP (1987), nicméně některé práce založené na centeringu tyto dva typy shiftu nerozlišují. Empirické zkoumání ale ukazuje, že tato distinkce je účelná, neboť přechod typu *rough-shift* je v přirozeném diskursu neexistující nebo extrémně vzácný.“<sup>40</sup> Tato neexistence, resp. velice vzácný výskyt typu

<sup>38</sup> Dále jsou užívány tyto zkratky:  $P_o$  – podmět,  $P_t$  – předmět,  $T$  – topik,  $F$  – fokus.

<sup>39</sup> Tento typ vztahu je využíván v případě první výpovědi kontextu, kdy není žádná předchozí výpověď, s níž by bylo provedeno srovnání.

<sup>40</sup> „The distinction between smooth and rough shift was first made by BFP, and some work in centering doesn't distinguish these two types of shifts. However, the distinction is supported by empirical

rough-shift je důvodem, proč s tímto typem v rámci zkoumání empirických dat v podstatě nepracujeme, i když je možné modelově takovýto typ přechodu vytvořit i v češtině – viz poznámka 52. Z toho zároveň vyplývá, že užití přechodu rough-shift se vyznačuje nejnižší mírou koherence textu.

Jako příklady jednotlivých typů přechodů poslouží následující modelové kontexty<sup>41</sup>. V modelových kontextech jsou pro zřetelnost navazování koreferenčních vztahů v textu uváděny plné jmenné fráze, v češtině by bylo přirozenější některé z nich pronominalizovat nebo elidovat. Pronominalizace ani elize by neovlivnily typ přechodu. První výpověď slouží jako výpověď předchozí ( $U_{i-1}$ , resp.  $U_{i-2}$ ), druhá a třetí jako výpověď aktuální ( $U_i$ ). V následujících příkladech u každé výpovědi uvádíme v tabulce D rozbor z hlediska distribuce center pozornosti ve výpovědi se zřetelem k výpovědi předcházející.

a. *Jeff pomáhal Dickovi mýt auto.*

Cp	[Jeff]
Cb	[?]
Cf	[Jeff, Dick, auto]
přechod	0

**Tabulka D1: Centering – analýza**

*Jeff myl okna, zatímco Dick voskoval.*

Cp	[Jeff]
Cb	[Jeff]
Cf	[Jeff, okna, Dick, auto]
přechod	continue

**Tabulka D2: Centering – analýza**

*Jeff vyleštil přední sklo.*

Cp	[Jeff]
Cb	[Jeff]
Cf	[Jeff, přední sklo]
přechod	continue

**Tabulka D3: Centering – analýza**

---

work that shows that rough-shift transitions are nonexistent or extremely rare in naturally occurring discourse (Di Eudenio, this volume), (Hurewitz, this volume).“ (WJP 1998: 6)

<sup>41</sup> Na základě WJP 1998: 6n.

b. *Jeff pomáhal Dickovi mýt auto.*

Cp	[Jeff]
Cb	[?]
Cf	[Jeff, Dick, auto]
přechod	0

**Tabulka D4: Centering – analýza**

*Jeff myl okna, zatímco Dick voskoval.*

Cp	[Jeff]
Cb	[Jeff]
Cf	[Jeff, okna, Dick]
přechod	continue

**Tabulka D5: Centering – analýza**

*Dick taky leštil kapotu.*<sup>42</sup>

Cp	[Dick]
Cb	[Dick]
Cf	[Dick, kapota]
přechod	smooth-shift

**Tabulka D6: Centering – analýza**

c. *Tom si koupil nový počítač.*

Cp	[Tom]
Cb	[?]
Cf	[Tom, nový počítač]
přechod	0

**Tabulka D7: Centering – analýza**

*Tom počítač ukázal Johnovi.*

Cp	[Tom]
Cb	[Tom]
Cf	[Tom, počítač, John]
přechod	continue

**Tabulka D8: Centering – analýza**

<sup>42</sup> Typ rough-shift by vypadal podobně, třetí věta v pořadí by zněla zhruba takto: *Dickem byla (vy)leštěna kapota., Kapota byla (vy)leštěna (Dickem).* V takovém případě by centra pozornosti byla rozložena takto: Cp [kapota], Cb [Dick], Cf [Dick, kapota], přechod: rough-shift. Tomuto typu přechodu se nevěnujeme, neboť se prakticky nevyskytuje. I v češtině je takové uspořádání diskurzu nepřirozené a nekompletní (informace o tom, že *Dickem byla vyleštěna kapota.* vyžaduje další pokračování výpovědi, např. *Dickem byla vyleštěna kapota, zatímco okna myl Jeff.*).



*John Tomovi vysvětlil jeho vylepšené funkce.*

Cp	[John]
Cb	[Tom]
Cf	[John, Tom]
přechod	retain

**Tabulka D9: Centering – analýza**

Příklad (a) představuje přechod typu continue, kdy platí vztah  $Cb = Cp$  (*Jeff*) a zároveň zůstává zachováno Cb, tj. platí i vztah  $Cb(U_i) = Cb(U_{i-1})$ . Příklad (b) je ukázkou přechodu typu smooth-shift, kde dochází ke změně Cb a zároveň je toto Cb i Cp aktuální výpovědi. Typ retain, tedy situaci, kterou bychom mohli popsat jako  $Cb \neq Cp$  ( $Tom_{Cb}$ ,  $John_{Cp}$ ) a zároveň  $Cb(U_{i-1}) = Cb(U_i)$ , představuje příklad (c).

Uvedené typy přechodů mezi Cp, Cb a Cf nejsou na první pohled vyčerpány, například chybí explicitní zkoumání vztahu mezi  $Cp(U_i)$  a  $Cp(U_{i-1})$ . Pokud ovšem prozkoumáme všechny důsledky a omezení, která vyplývají ze vztahů užívaných ke klasifikaci přechodů, tj. vztah mezi Cb a Cp aktuální výpovědi a vztahy mezi Cb výpovědi aktuální a předchozí, zjistíme, že vztahy mezi  $Cp(U_i)$  a  $Cp(U_{i-1})$  jsou jimi také postiženy a další kritérium by do typologie přechodů nepřineslo žádné obohacení.

Jednotlivé typy přechodů se v rámci výstavby souvislého textu kombinují, stejně jako v případě tematických posloupností by bylo velice neobvyklé, kdyby byl delší text vystavěn rekurzivně, tj. na základě pouze jednoho stále se opakujícího typu přechodu.

### 3.4.3 Zájmena

Zaměřujeme se na navazování takových koreferenčních řetězců v textu, v nichž jako anafor funguje zájmeno. Za základní zdroje informací o tom, jak se zájmeny zachází teorie centeringu, považujeme již zmíněné texty BFP (1987) a Candace Sidnerové (1981). Na tomto místě se zaměříme na přístup k zájmenům z hlediska centeringové teorie. Jedná se o již zmíněný algoritmus vázání zájmen (*algorithm for pronoun binding*),<sup>43</sup> který dělíme do

<sup>43</sup> Celý algoritmus je detailně představen v BFP (1987: 159) včetně příkladů, které proces přiřazování center pozornosti k entitám osvětlí (viz příloha Algoritmus a příklad, 10.1), na tomto místě příklad ani detailnější popis postupu analýzy z prostorových důvodů neuvádím.

tří fází: „Nejprve jsou možné navržené vazby<sup>44</sup> *konstruovány*, poté jsou *filtrovány* a nakonec *klasifikovány a hierarchizovány*.“<sup>45</sup>

V rámci první fáze jsou v aktuální výpovědi nalezeny všechny odkazující výrazy (*referring expressions RE's*), které jsou seřazeny podle gramatické funkce (1. Po, 2. Pt, 3. ostatní větné členy). Na tomto základě je vytvořen seznam možných Cf, v němž jsou zachyceny všechny Re rozšířené o informaci, zda jsou vyjádřeny jménem nebo zájmenem. Poté je vytvořen seznam možných Cb na základě předchozí výpovědi, který je doplněn o prázdnou hodnotu, aby byla zajištěna možnost, že ve výpovědi nebude žádné Cb. Následně jsou prostřednictvím zkombinování seznamu možných Cf a Cb vytvořeny tzv. navržené vazby, tedy všechny variance, potenciální kombinace, usouvztažnění entit a center pozornosti.

Ve druhé fázi je prováděn výběr ze všech vytvořených (logicky možných) navržených vazeb, a to aplikací tří filtrů. Prvním filtrem je tzv. kontraindexace (*contraindexing, contraindiced pronouns*): pokud je pro dvě kontraindexovaná zájmena indikován stejný antecedent, nebo pokud je indikován antecedent pro zájmeno, s nímž je kontraindexováno, je taková navržená vazba vyloučena. Typickým příkladem, kdy o vyloučení navržené vazby rozhodne kontraindexace, je případ sloves kontroly. Uspořádání výrazů do koreferenčního řetězce rozhoduje o tom, které výrazy budou kontraspecifikovány, tedy člen valenčního rámce predikátu a člen valenčního rámce slovesa kontrolovaného: *Petr poslal Pavla<sub>i</sub> [PRO]<sub>i</sub> běhat*. Druhým filtrem je srovnání seznamu možných Cf v rámci navržených vazeb se souborem Cf předchozí výpovědi (tento soubor musí být zachován v původním pořadí center pozornosti Cf). Cf předchozí výpovědi jsou postupně procházena a ta, která se vyskytují i v seznamu potenciálních Cf v navržených vazbách jsou zaznamenávána. Pokud entita potenciálního Cb v dané navržené vazbě neodpovídá prvnímu elementu tohoto seznamu, je taková navrhovaná vazba vyřazena. Pokud ke shodě mezi navrhovaným Cb a prvním elementem seznamu nedojde, je aplikován i třetí filtr, který eliminuje všechny navržené vazby, v nichž

<sup>44</sup> *proposed anchors* = doslova navržené kotvy, vazby; jedná se o všechny logické možnosti přiřazení Cb a Cf hodnot jednotlivým entitám výpovědi

<sup>45</sup> „First the proposed anchors are *constructed*, then they are *filtered*, and finally, they are *classified and ranked*.“ (BFP 1987) Následující popis algoritmu je volnou parafrází BFP 1987.

navrhované Cb neodpovídá Cf předchozí výpovědi realizovanému formou zájmena. Pokud v seznamu Cf předchozí výpovědi nejsou žádná Cf vyjádřena zájmenem, zůstává tento filtr nepoužit.

Třetí fáze spočívá v určování typu přechodu (continue, retain, smooth-shift, rough-shift). Každá navrhovaná vazba je klasifikována z hlediska typu přechodu, k němuž náleží. Na základě hierarchie typu přechodů představené výše je vytvořena hierarchie navrhovaných vazeb.

#### 3.4.4 Hypotézy o navazování koreferenčních řetězců

Teorie centeringu, na rozdíl od dosavadní české danešovské teorie, poskytuje predikce o tom, který typ výstavby textu, tedy přechodu v centeringové teorii, je nejčastější. V rámci úvah o četnosti jednotlivých typů přechodů teorie centeringu považuje za hlavní a nejčastější způsob výstavby textu typ přechodu continue, pro který argumentuje mimo jiné nejsnadnější kognitivní zpracovatelností.<sup>46</sup>

Typ přechodu continue bychom mohli zjednodušeně chápat jako navazování koreferenčního<sup>47</sup> řetězce na podmět předchozí věty, Cp(U<sub>i-1</sub>), prostřednictvím anaforu Cb(U<sub>i</sub>). Vycházíme z výše uvedené zásady (3.) o vztahu těchto dvou center pozornosti. Centering tedy poskytuje základ hypotézy o preferenci navazování na podmět.

Tato preference ovšem má několik protipříkladů: některé jazyky naznačují, že kritérium navazování na T má před kritériem navazováním na Po přednost: „jazyky, které poskytují identifikaci topikové funkce (např. japonština) naznačují, že topik je důležitější

<sup>46</sup> Tato preference spolu s hierarchií v rámci souboru Cf ovlivňuje strukturu výpovědi: „Strukturní paralelismus je důsledkem hierarchizace souboru Cf na základě gramatických funkcí a preference přechodu continue před retain.“ – „[S]tructural parallelism is a consequence of our ordering the Cf list by grammatical function and the preference for continuing over retaining.“ (BFP 1987: 157)

Za strukturní paralelismus je Kameyamovou (1986) označován vztah mezi morfosyntaktickou charakteristikou (zájmeno vs. ne-zájmeno) a větněčlenskou platností (podmět vs. ne-podmět): „subject pronouns prefer subject antecedents and non-subject pronouns prefer non-subject antecedents“ (BFP 1987: 157).

<sup>47</sup> Teorie centeringu se nezaměřuje pouze na koreferenci ve smyslu referenční totožnosti mimojazykových entit, k nimž výrazy – členy koreferenčního řetězce odkazují, nicméně zjednodušeně můžeme Cb považovat za anafor a centra pozornosti v předchozí výpovědi za potenciální antecedenty.

než podmět“<sup>48</sup>. V dotazníkovém šetření se tomuto aspektu věnujeme více na základě utváření koreferenčních vazeb v českém textu (viz oddíl Dotazníkový výzkum, 4.2).

Dalším „protipříkladem“ může být kontext typu *Brennan jezdí Alfou Romeo. Brennan jezdí rychle. Friedman s ní o víkendech závodí. Ona jede do Laguna Seca.*, kde podle BFP 1978, kteří tuto situaci zkoumali, „někteří respondenti preferovali spíše typ přechodu shift, přestože centeringový algoritmus predikuje spíše typ continue“<sup>49</sup>. Tato situace, tedy rozdíl mezi předpokládaným typem přechodu a typem přechodu interpretovaným respondenty, je variací situace zkoumané námi v rámci dotazníkového výzkumu, v němž jsou dotazníkové položky sestavovány tak, aby variovalo postavení jmenných frází v T a F, resp. v Po a Pt. Prostřednictvím takového nastavení výzkumu se ukáže, zda předpokládaný typ přechodu bude dominantním způsobem navazování koreferenčních řetězců.

---

<sup>48</sup> „languages that provide an identifiable topic function (e.g. Japanese) suggest that topic takes precedence over subjekt“ (BFP 1987: 156)

<sup>49</sup> „some informants seem to have a preference for shifting, whereas the centering algorithm chooses a continuation“ (BFP 1987:1 60)

## **4 Antecedent zájmena třetí osoby**

Nyní přejdeme k praktické části výzkumu. Dosavadní představené přístupy a teorie poskytují aparát, jakým zacházíme s jazykovými daty. Nejprve představíme zkoumání provedené na korpusových datech (Pražský závislostní korpus – PDT), poté dotazníkový výzkum, který byl proveden na vzorku necelých 300 rodilých mluvčích češtiny.

Z hlediska výstavby korpusů by byla korpusová metoda zkoumání navazování koreferenčních řetězců vhodnější spíše pro zkoumání produkční stránky textu, i když ani při takovémto zaměření výzkumu není bezproblémová. Vzhledem k tomu, že korpusy spravované Ústavem Českého národního korpusu jsou morfologicky otagované nejprve automaticky, posléze je provedena korekce anotátory a ne samotnými autory textů (z praktických důvodů), jde v podstatě o recepci stránku textu i přesto, že ÚČNK deklaruje zaměření korpusů na produkční stránku. Toto tagování by přesto mohlo být pro interpretaci výstavby textu přínosné, zejména při využití kongruenčních kategorií. Nicméně k takovému zkoumání nejsou korpusy ČNK uzpůsobeny, neexistují totiž vhodné nástroje, kterými by bylo možné zpracovat korpusové texty do takové podoby, aby koreference byla patrná. Syntakticky anotované korpusy (jako je například výše popsany Prague Dependency Treebank) jsou z tohoto hlediska jako zdroj textů vhodnější.

### **4.1 PDT**

Pražský závislostní korpus (*Prague Dependency Treebank* – PDT) je cca milionový korpus obsahující zhruba 50 000 vět anotovaných na třech rovinách – morfologické, analytické a tektogramatické. Základní přehled o PDT z hlediska aktuálního členění a koreferenčních vztahů poskytuje např. Kučová et al. 2005.

Pražský závislostní korpus v naší práci slouží jako zdroj referenčních dat. Byl zvolen z toho důvodu, že obsahuje oannotované syntaktické vztahy, na rozdíl od korpusů Českého národního korpusu. Anotace koreference a aktuálního členění probíhají v rámci tektogramatické roviny, lze je snadno propojit a získat tak data stejného typu, jaký je základem dotazníkového šetření.

#### **4.1.1 Aktuální členění v PDT**

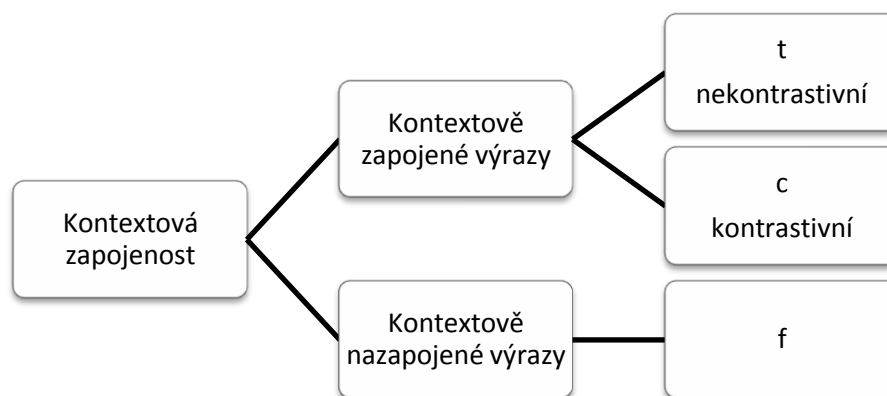
Značení aktuálního členění v PDT vychází ze dvou aspektů, které byly představeny v první části této práce – kontextové zapojenosti a výpovědní dynamičnosti. Tyto dva aspekty jsou

klíčové jednak pro řazení uzlů na tektogramatické rovině (kritérium výpovědní dynamičnosti), jednak pro určení hodnoty atributu *tfa*.

Na základě kontextové zapojenosti jsou částem výpovědi přiřazovány hodnoty atributu *tfa*<sup>50</sup>. Atribut *tfa* nabývá tří hodnot:

- hodnota *t* pro nekontrastivní kontextově zapojený výraz,
- hodnota *f* pro kontextově nezapojený výraz,
- hodnota *c* pro kontrastivní kontextově zapojený výraz.

Z přehledu hodnot, jichž nabývá atribut *tfa*, je zřejmé, že kromě kontextové zapojenosti je v rámci PDT zohledňováno i kritérium kontrastivnosti, které se uplatňuje až sekundárně a pouze v rámci kontextově zapojených částí textu. Jejich kombinaci zachycuje následující schéma I.



**Schéma I: anotace AČV v PDT**

Jednotlivé typy výrazů jsou anotovány na základě následujících pravidel:

- Za výraz typu *f* (*tfa* = *f*) je považován „výraz [...], který reprezentuje v textu dosud neznámá, nová fakta, nebo uvádí známá fakta do nových souvislostí, tj. vyjadřuje informace, které nejsou odvoditelné z kontextu“ (Mikulová et al. 2005: 143)
- Za výraz typu *t* (*tfa* = *t*) je považován „výraz [...], který uvádí do textu 'známou informaci'. Jde o výraz, který se opakuje (ne nutně doslova) z předcházejícího textu, nebo je z textu nějakým způsobem vyvoditelný (například pomocí koreferenčních či inferenčních vztahů), nebo jde o jednotku nějak související s širším kontextem.“ (Mikulová et al. 2005: 142n)

<sup>50</sup>

Tyto hodnoty jsou ve stromě zaznamenány světle modrou barvou vedle označení sémantické role.

- Za výraz typu c ( $tfa = c$ ) je považován „výraz, který můžeme identifikovat na základě těchto vlastností: a) výraz je zpravidla určen výběrem z množiny alternativ [...], b) [...] kontrastivní kontextově zapojený výraz se obvykle vyskytuje v jednotlivých oddílech výčtových pasáží, na začátcích odstavců apod., c) v mluvené podobě nese [...] fakultativní kontrastivní přízvuk.“ (Mikulová et al. 2005: 143)
- Jen typy t a f mohou být povrchově nevyjádřeny, typ c musí vždy být povrchově realizován.

Z uvedených pravidel přiřazování hodnot atributu  $tfa$  jednotlivým výrazům výpovědi je patrné, že není možné ztotožňovat topik (T) s výrazy kontextově zapojenými ( $tfa = t$ , c) a fokus (F) s výrazy kontextově nezapojenými ( $tfa = f$ ). Za modelové příklady považuje Mikulová et al. (2005: 143) případy typu: (*Dnes nesu jen dopis.*) *Tu knihu* <sub>$tfa=c$</sub>  *ti přinesu* *zítra* <sub>$tfa=f$</sub> . *Janu* <sub>$tfa=c$</sub>  *Marie* <sub>$tfa=t$</sub>  *neviděla* <sub>$tfa=f$</sub> .

#### 4.1.2 Koreference v PDT

Koreferenční vztahy v PDT jsou založeny na stejných termínech, jaké jsem představila výše (např. endofora, anaforické odkazování, antecedent, kataforické odkazování, postcedent). Existuje několik typů koreferenčních vztahů, které jsou v PDT anotovány, tj. koreference gramatická, koreference textová a tzv. bridging anafora. Ve verzi PDT 2.0 jsou anotovány první dva typy, i když z textové koreference pouze koreference zájmenná; jmenná textová koreference a bridging anafora jsou právě v procesu anotování.

V PDT je důsledně rozlišováno mezi koreferencí, tedy referenční totožností, a vztahy anaforickými, které se vyznačují kontextovou zapojeností, ale ne referenční totožností: „Koreference je reference dvou výrazů k témuž objektu reálného světa, zatímco anafora je odkaz na to, co bylo již v předchozím kontextu zmíněno.“ (Nědolužko 2011: 69) Z tohoto popisu je zřejmé, že anafora je pojem širší, „odkaz na to, co již bylo v předchozím kontextu zmíněno“ zahrnuje i referenci k objektu reálného světa. Anaforické vztahy, v nichž nedochází ke koreferenci, jsou anotovány jako asociační anafora (vztahy založené na inferenci mezi nominálními vyjádřeními, tzv. bridging vztah) několika typů, které

uvádíme i s příklady bez dalšího komentáře, přičemž výrazy ve vztahu asociační anafory jsou podtrženy:

- část – celek, např. *Trojnásobný mistr světa a autor tohoto výroku nepřežil VC San Marina, když v třísetkilometrové rychlosti nevytočil zátáčku Tabburello a narazil do betonové zdi. Po sobotní smrtelné nehodě Rakušana Rolanda Ratzenbergera si okruh v Imole připsal druhý křížek a v Brazílii byl vyhlášen třídení státní smutek.*
- množina – podmnožina/prvek, např. *Pavel Vondráček: Termín převýchova znám pouze z nacistického a komunistického slovníku. Na převýchovu se, pokud vím, posílali ti, kteří měli podle těchto zruďných režimů nevhodný původ.*
- entita – entita s jedinečnou funkcí v jejím rámci, např. *Na přímou podporu podnikání vydá letos stát přibližně 1,8 procenta hrubého domácího produktu. Tuto skutečnost jednoznačně konstatuje ministr hospodářství Karel Dyba v analýze, kterou předložil vládě.*
- kontrast sémantického a kontextového protikladu, např. *Dnes, po rozdělení ČSFR, je jasné, že osud ČR bude stále více spojený s Německem a přes něj s Evropskou unií a osud Slovenska s Ruskem.*
- anaforické odkazování mezi nekoreferenčními entitami, např. *Rozbití Varšavské smlouvy bylo jako odseknutí údů od těla. Od té doby se toho mnoho neudělalo.*

Anotace koreference v PDT probíhá jednofázově<sup>51</sup> a je prováděna na dvou úrovních, první z nich představuje koreference gramatická, druhou koreference textová. Rozdíl mezi nimi vyplývá ze způsobu, jakým jsou koreferované členy řazeny do koreferenčních řetězců. „Gramatická koreference je taková koreference, kdy je možné určit koreferovaný člen na základě gramatických pravidel.“ (Mikulová et al. 2005: 129) Oproti tomu „[t]extová koreference je běžně chápána jako užití různých jazykových prostředků [...], které většinou anaforicky odkazují. Toto odkazování není realizováno pouze gramatickými prostředky, ale i na základě kontextu.“ (Mikulová et al. 2005: 139)

<sup>51</sup> „na rozdíl od většiny schémat anotace koreference a anaforických vztahů nedělíme proces anotace do dvou fází: vyhledávání potenciálních anaforických jednotek, kandidátů na anaforický vztah (tzv. 'markables'), a anotaci komplexního vztahu“ (Nědolužko 2011: 73)



Typickým příkladem gramatické koreference jsou zvrtná zájmena (*Pavla se prohlíží v zrcadle.*), vztažné prostředky (*Ti, co kroutí hlavami...*, *Mužstvo získalo tři body, což je maximum.*), doplněk nebo jemu analogická pozice (*Odcházela poražena.*, *Mužstvo zůstává neporaženo.*) nebo kontrola (*Podnik plánoval zvýšit výrobu.*). Textová koreference nemá tak jednoznačně vymezené prostředky, anotace se týká všech jmenných frází, některých adjektivních frází a některých adverbii. Textová koreference je dále tříděna podle rozsahu antecedentu na explicitní koreferovaný člen, koreferovaný segment (větší úsek textu) a koreferenci exoforickou.

Předmětem našeho výzkumu je pouze koreference textová. V rámci textové koreference existuje několik typů odkazování, které jsou vymezeny různou mírou specifičnosti reference a konkrétnosti významu.<sup>52</sup> První skupinu tvoří fráze referenční, které odkazují k prvkům reality a mohou sloužit jako antecedenty, druhou skupinu tvoří fráze nereferenční, které nejsou anotovány, podle Mendozové „nemohou vystupovat jako antecedenty anaforického vztahu“ (Nědolužko 2011: 28).

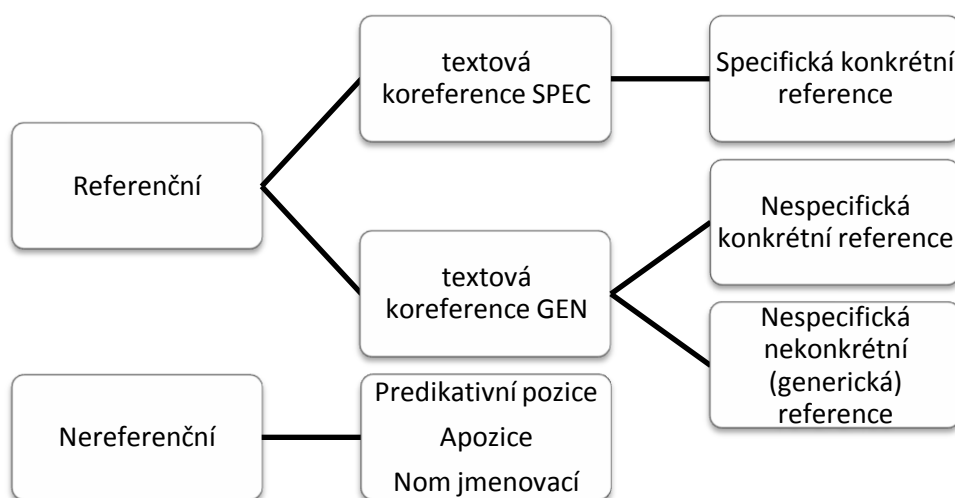


Schéma J: Specifičnost reference

V rámci skupiny referenčních jmenných frází jsou v PDT vydělovány tři podskupiny: jmenné fráze s konkrétní specifickou referencí s atributem SPEC (např. *Helena prosila maminku<sub>i</sub>, aby na ni počkala. Maminka<sub>i</sub> však řekla, že nemůže.*), s referencí

<sup>52</sup> Uvedená typologie specifičnosti reference v PDT vychází z Mendozové (2004), zpracováno podle Někdužko (2011).

nespecifickou, ale stále konkrétní, tedy fráze, v nichž existuje konkrétní objekt, k němuž je odkazováno, nicméně není specifikován (atribut GEN, např. *Poprosím o to [některého z kolegů]<sub>i</sub>; a ten<sub>i</sub> mi to řekne.*) a s referencí nespecifickou, a tedy nekonkrétní, tj. generická reference (atribut GEN, např. *Děti milují čokoládu.*), která odkazuje k prototypickému představiteli entity. Při srovnání jednotlivých typů referenčních vztahů vyvstává otázka, zda generická užití považovat za typ koreferenčních vztahů, když nemůžeme nikdy zaručit referenční totožnost. V PDT se přiklání k názoru, že ano, neboť se podílí na koherenci textu a „generické jmenné fráze se mohou účastnit veškerých anaforických vztahů a podobně jako jmenné fráze se specifickou referencí mohou být pronominalizovány“ (Nědolužko 2011: 100).

Nereferenční jmenné fráze nejsou v PDT v rámci koreference anotovány. Jedná se o tyto případy: predikátové užití typu *Petr<sub>i</sub> je programátor<sub>i</sub>.*; apozice (*Petr<sub>i</sub>, náš programátor<sub>i</sub>, má zítra státnice.*) a užití nominativu jmenovacího (*Podle těchto zpráv nějaká firma<sub>i</sub> na naše území umísťuje německou delikventní mládež. V Košťanech totiž zakoupila dům firma<sub>i</sub> Struktura<sub>i</sub>, která se u nás rozmisťováním německých chlapců zabývá.*).

Na takto oannotovaných datech jsme provedli měření počtu koreferenčních vztahů s ohledem na aktuálně- a větněčlenskou platnost koreferovaných výrazů. Postup a výsledky tohoto měření představují následující pasáže textu.

#### 4.1.3 Zkoumání koreference

Koreferenci jsme zkoumali na všech datech PDT verze 2.0, přičemž v úvahu jsou brány jen textové koreference. Důraz je kladen na koreferenci textovou, protože gramatická koreference neodpovídá základnímu předpokladu pro realizaci dotazníkového výzkumu, tedy kontextu, v němž jsou koreferenční vztahy nejednoznačné (viz dále), v rámci gramatické koreference a) jsou koreferenční vztahy dané, konkrétně gramatickými pravidly daného jazyka, b) jsou zahrnuty i koreferenční vztahy v textu reprezentované pomocí jiných typů zájmen, než na které se zaměřujeme v dotazníkovém výzkumu. Vztahy zahrnuté do měření zachycují textovou koreferenci mezi členy nebo částmi bezprostředně sousedících výpovědí, nezabýváme se tedy koreferenčními textovými vztahy, které nejsou realizovány přes hranici věty. Východiskem zkoumání je koreference textová i z toho

důvodu, že v rámci gramatické koreference jsou anotovány i vztahy, které nejsou předmětem našeho zkoumání (viz oddíl Zájmenná koreference, 4.2.5). Nevyhovující vztahy filtrujeme následovně: Anaforem ve zkoumaném koreferenčním vztahu textové koreference je vždy zájmeno a antecedentem jmenná fráze v předchozí výpovědi.

Zároveň jde pouze o vztahy anaforické v tom smyslu, že odkazují k předcházejícímu kontextu, členy koreferenčního řetězce se v textu vyskytují v tomto pořadí: 1. antecedent v první výpovědi, 2. anafor ve výpovědi bezprostředně následující. V rámci zkoumání koreference v PDT, které slouží jako srovnávací data pro dotazníkový výzkum, se řídíme několika následujícími omezeními.

Aby byla zajištěna co největší podobnost s kontexty v dotazníkovém výzkumu, byly měření podrobeny kontexty s efektivním<sup>53</sup> bratrským uzlem<sup>54</sup>, tedy kontexty, v nichž je ve výpovědi s antecedentem více jmenných frází, které mohou jako antecedent sloužit, např. typu *Ekonomická analýza potvrzuje správnost nově vytvořené strategie. Sklízí pochvaly.*

Shrňme tedy, že v rámci PDT jsou zkoumány vztahy textové koreference, které jsou realizovány přes hranici věty tak, že anafor (zájmeno) je součástí druhé výpovědi v pořadí a jako jeho antecedent může sloužit více jmenných frází v předchozí výpovědi. Gramatická koreference a bridgingové vztahy (asociační anafora) nejsou předmětem měření. Vyhledávání vztahů podle zadání provedl Jiří Mírovský „programátorským jazykem“.<sup>55</sup> Textové koreferenci<sup>56</sup> v rámci vymezených kritérií odpovídá 6 126 koreferenčních vztahů.

<sup>53</sup> Za efektivní uzly jsou v technické dokumentaci PDT označovány ty uzly, které jsou ve vztahu závislosti: první, řídící, uzel je označován jako efektivní rodič, druhý, závislý, uzel jako efektivní potomek. Pokud existuje více uzlů závislých na stejném efektivním rodiči, jsou tyto uzly ve vzájemném vztahu označovány jako efektivní sourozenecké uzly.

<sup>54</sup> V PDT je tradičně používán termín sestra uzlu, tedy sesterský uzel. My jsme se ale rozhodli používat termín bratr, bratrský uzel, protože jsme chtěli dát jasně najevo, že jde o specifický typ vztahu mezi uzly.

<sup>55</sup> dotazy byly sestaveny (s různými obměnami) takto: t-root \$n := [ descendant t-node, [ gram/sempos ~ "^v", echild t-node \$n4 := [ gram/sempos ~ "^n", a/lex.rf a-node \$n7 := [ ] ], echild t-node \$n3 := [ gram/sempos ~ "^n", a/lex.rf a-node \$n6 := [ ] ] ], same-document-as t-root [ tree\_ord = \$n.tree\_ord+1, descendant t-node \$n2 := [ gram/sempos ~ "^n.pron", coref\_text/target-node.rf \$n3, a/lex.rf a-node \$n8 := [ ]

#### 4.1.4 Výsledky

Textová koreference mezi zájmenem a jmenným antecedentem na větší vzdálenost, než je bezprostředně předcházející výpověď, je zastoupena pouze 442 vztahy. (Pokud nebudeme omezovat typ antecedentu na jmennou frázi, je takových vztahů 532.) V rámci všech vztahů mezi zájmenem a jmennou frází jako jeho antecedentem distanční navazování koreferenčních řetězců představuje jen zlomek (13.86 %). Je tedy zřejmé, že textová koreference je v češtině mezi jmennou frází a zájmenem, které ji zastupuje (na základě jazykového materiálu PDT, tedy převážně v publicistických textech, a za výše uvedených omezení), realizována především v kontaktním postavení výpovědi.

Koreferenční vztahy můžeme na základě hodnot atributu tfa, kterých nabývají členy koreferenčních řetězců, rozčlenit. Pro naše potřeby nerozlišujeme kontrastivní (tfa = c) a nektrastivní (tfa = t) kontextovou zapojenost, protože klíčové je rozlišení na kontextově zapojené (tfa = t, c) a kontextově nezapojené prvky (tfa = f). Následující tabulka zachycuje data získaná vyhledáním všech koreferenčních řetězců mezi bezprostředně sousedícími výpověďmi (tabulka E).

	do t/c	do f	sum
z t/c	4 515	1 565	6 080
z f	31	15	46
sum	4 546	1 580	6 126

**Tabulka E: textová koreference (tfa)**

Anaforem v koreferenčním řetězci se v případě textové koreference stávají nejčastěji zájmena kontextově zapojená, poměr vůči anaforům v kontextově nezapojeném postavení je 2.88 a tento rozdíl je signifikantní ( $X^2$  test: chi-squared = 132.17, df = 1, p-value < 2.2e-16)<sup>57</sup>. Jde o kontexty typu *Jeho dojetí znásobila při vyhlásování přítomnost*

---

```
] ] ]; >> give $n3.id,$n3.tfa,$n2.tfa,$n6.afun,$n7.afun,$n8.afun >> for $1,$2,$3,$4,$5,$6 give distinct $1,$2,$3,$4,$5,$6 >> for $2,$3,$4,$5,$6 give $1,$2,$3,$4,$5,count()
```

<sup>56</sup> Dále pracujeme s pojmem textová koreference, aniž bychom uváděli, čím je omezena (viz výše uvedená kritéria vymezující textově koreferenční vztahy, které byly předmětem měření), nicméně je nutné mít tato omezení na paměti.

<sup>57</sup> Ověřování pravděpodobnosti chyby bylo provedeno na základě testu zvaného  $X^2$  (chí-kvadrát). Parametry, které uvádíme, představují naměřenou hodnotu  $X^2$ , počet stupňů volnosti a hodnotu

pořadatelů soutěže Českého manažerského centra v Čelákovících. Toho<sub>tfa=t</sub> se letos zúčastnilo 30 soutěžících. Koreferenční vztahy „z f“, tedy z kontextově nezapojeného anaforu ve druhé větě, tvoří jen minimum z celkového množství koreferenčních vztahů, což je způsobeno zejména tím, že postavení v rámci skupiny kontextově nezapojených výrazů (ve smyslu nevývoditelnosti z kontextu) je zdánlivě neslučitelné s účastí v koreferenčním vztahu (ve smyslu referenční totožnosti s entitou vystupující v předcházející části textu). Přesto jsou ale takové koreferenční vztahy možné, příkladem by mohly být následující situace: *Petr<sub>i</sub> běžel lesem. Tohle jsou jeho<sub>i</sub> špinavé kalhoty.*<sup>58</sup> (koreferenční vztah z f do t), nebo situace *Byl tady jen Petr<sub>i</sub>. To byl on<sub>(i)</sub>, kdo<sub>i</sub> snědl všechny buchty.* (koreferenční vztah z f do f). Je však zřejmé, že koreferenční vazby mezi kontextově nezapojeným výrazem druhé výpovědi a kontextově zapojeným nebo nezapojeným výrazem výpovědi první jsou velice řídké.

Koreferenční vazby z t/c do t/c jsou typickými případy Danešova průběžného tématu, vztahy z t/c do f naopak návazné tematizace rématu. To platí pro gramatickou i textovou koreferenci. Z dat PDT se dá usuzovat, že průběžné téma (navazování na kontextově zapojené výrazy) je v češtině častější než tematizace rématu, o které Daneš tvrdí, že je v češtině dominantním schématem tematické posloupnosti. Na první pohled je patrné, že poměr mezi navazováním na t a f v textové koreferenci je téměř trojnásobný (2.88).<sup>59</sup> Rozdíly mezi počtem koreferenčních vazeb do t/c (4977) a do f (2477) jsou signifikantní (prokázáno pomocí  $X^2$  testu:  $\chi^2 = 838.476$ ,  $df = 1$ ,  $p\text{-value} < 2.2e-16$ ), v celkovém součtu je tedy navazování na t/c signifikantně častější než navazování na f.

Data z PDT slouží především jako podklad pro dotazníkový výzkum, jako referenční materiál. Z toho důvodu se v rámci všech korpusových textů budeme soustředit jen na ta

---

pravděpodobnosti (p-value), která je pro určení statistické signifikantnosti výsledků klíčová. I v českém prostředí jsou standardně užívány statistické programy operující v angličtině, proto jsou všechny parametry zaznamenávány podle anglického úzu, včetně číselných hodnot (desetinná tečka).

<sup>58</sup> Odhlížíme od potenciální nejednoznačnosti kontextu v případě, že *Petr* a *jeho* nejsou koreferovanými členy jednoho koreferenčního řetězce.

<sup>59</sup> Hodnoty jsou standardně zaokrouhlovány na dvě desetinná místa s výjimkou hodnot pravděpodobnosti (p-values), které jsou zaokrouhlovány na tři desetinná místa.

data, která simulují situaci dotazníkového výzkumu, kdy v první větě jsou dvě jmenné fráze jako potenciální antecedenty zájmena z věty druhé. Celkové množství takových vztahů, v nichž existuje efektivní bratrský uzel, tedy potenciální konkurent při volbě antecedentu, představuje 2 064 vztahů. Je ovšem nutné říci, že těchto 2064 vztahů zahrnuje všechny typy bratrských jmenných frází, tedy jmenné fráze v postavení všech větných členů, v nichž má jmenná fráze možnost vystupovat a byly v PDT nalezeny. Rozložení bratrských uzlů v klíčových větněčlenských funkcích (Po – 523, Pt – 569) je rovnoměrné, mírně převažuje postavení v Pt (o 46 případů). Jmenné fráze jakožto potenciální antecedenty zastupují ve výpovědích funkce Po a Pt v cca polovině případů, druhou polovinu tvoří všechny ostatní větné členy (972). V případech, kdy efektivní bratrský uzel zastupuje některý z uvedených větných členů výpovědi, není samozřejmě vyloučeno, že zvolený antecedent plní stejnou větněčlenskou roli (např. několikanásobný podmět). Naopak, prototypickou situací výskytu efektivního bratrského uzlu je koordinací spojení. Z toho důvodu není větněčlenské postavení bratrského uzlu faktorem ovlivňujícím navazování koreferenčního řetězce.

Následující tabulky zachycují navazování koreferenčních řetězců s ohledem na existenci efektivního bratrského uzlu v první výpovědi, tedy v situaci více potenciálních antecedentů,<sup>60</sup> tabulka F z perspektivy aktuálněčlenské, tabulka G z perspektivy větněčlenské.

	do t/c	do f	sum
z t/c	1 178	853	2 031
z f	21	12	33
sum	1 199	865	2 064

**Tabulka F: koreference v ef. br. (AVČ)**

Z tabulky G je zřejmé, že v případě aktuálního členění jako anafor zcela dominuje zájmeno v t/c (2031:33, 61.55;  $X^2$  test: chi-squared = 1934,11, df = 1,  $p < 2.2e-16$ ), jako antecedent dominuje také jmenná fráze v t/c, i když ne tak výrazně (1199:865, 1.39). Dominance postavení t/c v případě anaforu je logická a očekávaná, protože (koreferovaný)

<sup>60</sup> Pro zjednodušení a přehlednost je situace, v níž dochází k výskytu efektivního bratrského uzlu první výpovědi, označována jako ef. br.

anafor je kontextově zapojený, tudíž je v neutrálním vyjádření je i jako kontextově zapojený, tj. v t/c, ztvárněn a podáván adresátovi. Dominantní postavení antecedentu v t/c i v případě, že v první výpovědi existuje efektivní bratrský uzel, potvrzuje naši hypotézu o preferenci navazování koreferenčních vazeb spíše na T než na F, pokud k problematice přistoupíme velice zjednodušeně a ztotožníme kontextově zapojené výrazy s T a kontextově nezapojené výrazy s F ( $X^2$  test: chi-squared = 54.05, df = 1, p-value = 1.956e-13). Výrazná převaha navazování koreferenčních řetězců z t/c do t/c svědčí o tom, že tematická posloupnost průběžného tématu je v českém materiálu převažujícím způsobem utváření koreferenčních vztahů a zajišťování koherence textu. Druhým nejčastějším typem je potom tematizace rématu. Toto pořadí odporuje Danešovu předpokladu, že tematizace rématu je častější metodou výstavby souvislého textu. Ke stejnému závěru jsme již dospěli v případě měření počtu koreferenčních vztahů mezi antecedentem bez efektivního bratrského uzlu. (Ne)přítomnost efektivního bratrského uzlu tedy podle tohoto měření neovlivňuje navazování koreferenčních řetězců v textu.

	do Po	do Pt	jínam	sum
z Po	164	219	97	480
z Pt	261	203	114	578
odjinud	517	286	203	1 984
sum	942	708	414	2 064

**Tabulka G: koreference v ef. br. (VČ)**

V případě větněčlenské platnosti<sup>61</sup> je v celkovém součtu anafor umístován spíše do postavení Pt než do Po (z Po: z Pt = 480:578, 0.83) a i v tomto případě je tato mírná převaha signifikantní ( $X^2$  test: chi-squared = 9.08, df = 1, p-value = 0.00). Zajímavé ovšem je, že tato tendence k navazování koreferenčního řetězce do předchozí výpovědi od Pt není konzistentní v tom smyslu, že v případě odkazování do Po předchozí výpovědi je Pt jako anafor častější, ale v případě odkazování na Pt předchozí výpovědi (tedy situace, kdy předmět první výpovědi zůstává předmětem i v druhé) je častější než postavení anaforu

<sup>61</sup> Zaměřujeme se pouze na postavení v Po a Pt, postavení ve funkci ostatních větných členů uvádíme jen pro úplnost, jsme si ho vědomi. V rámci výzkumu s ním nepracuji jako s hodnotou proměnné větněčlenské platnosti, nicméně koreferenční vazby vycházející z dalších větných členů do antecedentu v Po nebo Pt předchozí věty jsou započítávány do celkového počtu koreferenčních odkazů k Po, resp. Pt.

v Pt spíše postavení v Po (tedy situace, kdy se předmět první výpovědi stává podmětem druhé). Změna (Po – Pt) je tedy typičtější, než zachování linie Po – Po, resp. Pt – Pt. Nicméně tato distinkce nebyla na našich datech spolehlivě prokázána ( $p$ -value = 0.77). Naopak to, že do pozice podmětu je signifikantně častěji odkazováno z předmětu druhé výpovědi (0.63), je doložený fakt ( $X^2$  test:  $\chi^2$ -squared = 22.14,  $df$  = 1,  $p$ -value = 2.536e-06). Příkladem mohou být kontexty, v nichž se z předmětu navazuje na podmět, nebo kontexty, v nichž dochází k navazování z podmětu na předmět.

Antecedentem zájmena se stává spíše Po než Pt (942:708, 1.33), což potvrzuje hypotézu o častějším navazování na podmět ( $X^2$  test:  $\chi^2$ -squared = 33.19,  $df$  = 1,  $p$  = 8.377e-09).

Následující tabulka H poskytuje přehled toho, jak jsou data distribuována z hlediska obou kritérií, tedy aktuálního členění i větněčlenské platnosti. Závěry o navazování na t/c, resp. f, představené výše korespondují s daty zde uvedenými.

V kombinaci proměnných aktuálněčlenské a větněčlenské nacházíme na první pohled rozdílné výsledky v případě dat kontextově zapojených a kontextově nezapojených. Vztahy zahrnující výrazy kontextově zapojené (tj. vztahy z t/c a vztahy do t/c) jsou výrazně častější než vztahy zahrnující výrazy kontextově nezapojené (z f, do f). Distinkce větněčlenská tak výrazná není, nicméně vztahy zahrnující Po jsou častější než vztahy zahrnující Pt. Z tabulky H vidíme, že nejvíce koreferenčních vztahů (pokud necháme stranou ostatní větné členy) je navazováno mezi kontextově zapojeným předmětem a kontextově zapojeným podmětem předchozí výpovědi.

		do t/c 1 199			do f 865		
		661 do Po	333 do Pt	205 jinam	281 do Po	375 do Pt	209 jinam
<b>z t/c</b>	2031	653	325	200	280	369	204
z Po	468	95	66	32	67	145	63
z Pt	571	201	114	54	58	87	57
odjinud	992	357	145	114	155	137	84
<b>z f</b>	33	8	8	5	1	6	5
z Po	12	2	5	1	0	3	1
z Pt	7	2	2	3	0	0	0
odjinud	14	4	1	1	1	3	4

**Tabulka H: t/c, f + Po, Pt v ef. br.**



4.1.5 *Shrnutí*

Z výsledků měření dat týkajících se koreference v PDT (zachycených v tabulce I) vyplývá několik závěrů.

	vzájemný poměr	X <sup>2</sup>	p-values
textová koreference			
z t/c	6 080:46	5 943.40	p < 2.2e-16
z f	132.17		
do t/c	4 546:1 580	1 436.04	p < 2.2e-16
do f	2.88		
ef. br.			
z t/c	2 031:33	1 934.11	p < 2.2e-16
z f	61.55		
do t/c	1 199: 865	54.05	p = 1.956e-13
do f	1.39		
z Po	480:578	9.08	p = 0.003
z Pt	0.83		
do Po	942:708	33.19	p = 8.377e-09
do Pt	1.33		

**Tabulka I: přehled výsledků měření (PDT)**

1. Postavení antecedentu v rámci kontextově zapojených uzlů předcházející výpovědi je častější než postavení v rámci kontextově nezapojených uzlů. Toto zjištění potvrzuje hypotézu o preferenci navazování koreferenčních řetězců na kontextově zapojenou (zjednodušeně řečeno topikovou) část výpovědi, z hlediska přístupu centeringové teorie tedy existuje preference k utváření koreferenčních řetězců prostřednictvím typů výstavby textu continue a retain.
2. Zároveň jsme ukázali (tabulka F), že preference průběžného tématu je výrazně dominantním principem výstavby textu.<sup>62</sup>
3. Výběr antecedentu je v rámci měření textové koreference více příznakový než výběr aktuálněčlenského postavení anaforu, protože převaha postavení anaforů v t/c

<sup>62</sup> K potvrzení tohoto předpokladu není prosté konstatování o převažujícím způsobu navazování koreferenčních řetězců směrem k topiku předchozí výpovědi dostačující, neboť je potřeba vzít v úvahu také postavení anaforu zmíněného koreferenčního řetězce. Tento postup byl aplikován v tabulce J a příslušné pasáži této práce.

oproti jejich postavení v f je výraznější než převaha postavení antecedentů v t/c oproti postavení v f.

4. Postavení anaforu z hlediska větněčlenské platnosti se zdá být ustálenější než postavení antecedentu.
5. Anaforem koreferenčního vztahu je z hlediska větněčlenského nejčastěji předmět předchozí výpovědi.
6. V rámci dat simulujících dotazníkovou situaci (anaforem je zájmeno odkazující do předchozí výpovědi k jedné ze jmenných frází) nejsou nalezené tendence tak přesvědčivé, kromě jasné preference umisťování anaforu do t/c, přestože jsou všechny výsledky signifikantní.
7. Existuje tendence preferovat navazování koreferenčního řetězce na podmět předchozí věty (tedy typ continue).
8. Reference k Po byla prokázána signifikantně častěji v případech, kdy anaforem koreferenčního vztahu byl ve druhé výpovědi Pt.

## 4.2 Dotazníkový výzkum

### 4.2.1 Cíl výzkumu

Cílem výzkumu je popsat způsoby navazování koreferenčních vztahů v textu. Jako prostředek k získání dat slouží dotazníkové šetření. Tento výzkum by měl ukázat, zda jsou predikce vycházející z typologie způsobů koreferenčního navazování v textu představené v kapitole Centering prakticky využitelné i pro český materiál. V tomto výzkumu se zaměřujeme na recepční stránku textu, tedy interpretaci možných koreferenčních vztahů v textu jeho příjemcem. Vedou nás k tomu dva vzájemně neoddělitelné důvody. Prvním z nich je fakt, že se jedná o první a mapovací výzkum, jehož cílem je zjistit, zda je takové testování realizovatelné, zda existují nějaké prokazatelné tendence projevující se v datech, zda má vypovídající hodnotu atd. Druhým důvodem je fakt, že pro takovýto typ výzkumu je zkoumání produkční stránky textu velice složité, bylo by k němu pravděpodobně zapotřebí testování takzvanými on-line metodami.

### 4.2.2 Metoda výzkumu

Existuje několik různých metod, které se pro lingvistický výzkum v současnosti nabízejí. První z nich je korpusový výzkum. Zkoumat, jakým způsobem jsou recipienty navazovány<sup>63</sup> vztahy v textu, ale čistě pomocí morfologicky otagovaného korpusu není vhodné, jak již bylo zmíněno výše.

Druhou metodou, která se ke zkoumání navazování koreferenčních vztahů v textu nabízí, je dotazníkové šetření. Tato metoda byla pro současné zkoumání zvolena především z toho důvodu, že větší počet respondentů eliminuje přílišný vliv jediné interpretace (jako by tomu bylo například u korpusových dat, kde je počet anotátorů omezený). Navíc poskytuje možnost přizpůsobit data přímo specifické testovací situaci.

Design dotazníkového šetření můžeme shrnout do dvou základních bodů:

---

<sup>63</sup> Při zaměření výzkumu na recepční stránku textu užívám v souvislosti s koreferenčními vztahy pojmy, jako jsou utváření, navazování apod. Všechny označují způsob, jakým recipient interpretuje koreferenční řetězce v textu, některé z nich ale akcentují skutečnost, že tato interpretace nemusí odpovídat autorem zamýšlené výstavbě vztahů v textu, tedy že k utváření koreferenčních řetězců dochází na straně recipienta nezávisle na autorovi a v některých případech i na místech, kde to autor nezamýšlel. Recipient tedy v jistém smyslu textové vztahy vytváří stejně jako autor.

1. Dotazník obsahuje celkem 50 položek, 16 testovacích a 34 výplňkových. Tento poměr dvou výplňkových položek ku jedné testovací byl zvolen na základě standardního nastavení, tedy požadavku minimálního poměru 1:2.
2. Testovací kontexty jsou uměle vytvořené. Problematice sestavování testovacích kontextů je věnována samostatná kapitola.

#### 4.2.3 *Proměnné*

Jak již bylo naznačeno, náš výzkum se zabývá dvěma navzájem nezávislými proměnnými. První z nich je aktuální členění a druhou větněčlenská platnost. Hypotézy s nimi spojené testují, zda postavení jmenné fráze v topiku nebo fokusu (resp. v podmětu nebo předmětu) ovlivní interpretaci koreferenčního řetězce. Pro úplnost uvádím i hypotézy týkající se obou proměnných, hypotéza A aktuálního členění, hypotéza B větněčlenské platnosti.

(A) Navazování koreferenčního řetězce probíhá častěji na entitu zastoupenou v pozici topiku než fokusu.

(B) Navazování koreferenčního řetězce probíhá častěji na entitu zastoupenou v pozici podmětu než předmětu.

Důležitým předpokladem pro zkoumání vlivu aktuálního členění a větněčlenské platnosti na navazování koreferenčních řetězců je vymezení jednotlivých hodnot. Vzhledem k výše zmíněné problematické terminologii a vymezování topiku a fokusu považujeme za postavení v topiku první pozici ve větě (v rozsahu fráze), za postavení ve fokusu pozici poslední. Modelovým příkladem může být například testovací položka *Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.* V následujících příkladech uvádím jednotlivé notace tak, aby bylo zřejmé, která část výpovědi je v našem přístupu považována za topik, fokus, podmět a předmět. Pro úplnost uvádím všechny varianty této testovací položky.

*[Galileo]T veřejně podporoval [Koperníka]F. Věřil, že Země se otáčí kolem Slunce.*

*[Koperníka]T veřejně podporoval [Galileo]F. Věřil, že Země se otáčí kolem Slunce.*

*[Galileo]Po veřejně podporoval [Koperníka]Pt. Věřil, že Země se otáčí kolem Slunce.*

*[Galileem]Pt byl veřejně podporován [Koperník]Po. Věřil, že Země se otáčí kolem Slunce.*

Z příkladů je zřejmé, že při zkoumání jedné proměnné se mění v rámci dvou variant vždy jen hodnoty této zkoumané proměnné, druhá proměnná zůstává z hlediska nabývání svých hodnot stejná, aby bylo vyloučeno ovlivnění výsledků druhou proměnnou. V prvním příkladu je zkoumanou proměnnou aktuální členění, tudíž se postavení jmenných frází *Galileo* a *Koperník* v topiku a fokusu proměňuje, ale jejich větněčlenská platnost zůstává zachována (*Galileo* – Po, *Koperník* – Pt). Naopak ve druhém příkladu je zkoumanou proměnnou větněčlenská platnost, jmenné fráze *Galileo* a *Koperník* nabývají střídavě hodnot Po a Pt, jejich postavení v T a F se ale nemění (*Galileo* – T, *Koperník* – F).

V rámci zkoumání druhé proměnné, tedy vlivu větněčlenské platnosti na navazování koreferenčních řetězců, se nabízí velké množství hodnot (větných členů), kterých proměnná může nabývat a nabývá (např. podmět, předmět, přívlastek, predikát, příslovečné určení). Za zástupce této proměnné, které budou předmětem zkoumání, byly zvoleny podmět a předmět z těchto důvodů:

- Podmět a předmět představují základní větné členy v tom smyslu, že jsou predikátem valenčně vázané ve větě (v případě předmětu máme na mysli především předmět přímý). Podmět je ve větě přítomen vždy, minimálně v rámci hloubkové struktury, tedy ne nutně povrchově, je obligatorní, pokud vyloučíme věty jednočlenné.
- Na rozdíl od nich jsou větné členy jako přívlastek nebo příslovečné určení spíše doplňkové v tom smyslu, že jejich přítomnost ve výpovědi nemá zásadní vliv na interpretaci jejího významu, nejsou gramatikalizované a ve většině případů ani hloubkově obligatorní.<sup>64</sup>

<sup>64</sup> Jde samozřejmě o generalizaci, jsme si plně vědomi toho, že například přívlastek může mít zásadní vliv na význam výpovědi, např. *Mladší děti musí být očkovány.* vs. *Starší děti musí být očkovány.* Je také

- Podmět a předmět jsou základními formami reprezentace sémantických rolí ve výpovědi. Typické uspořádání je agens v pozici podmětu, patient v pozici předmětu. Podmět je tedy levovalenčním doplněním, předmět doplněním pravovalenčním. V češtině je situace, kdy dochází k jejich záměně, při jiné hierarchizaci větné propozice (srovnejme *Naše babička upekla koláče. a Koláče byly upečeny babičkou.*), běžná.

Proměnné se překrývají, každý kontext vykazuje charakteristiky vymezené z hlediska aktuálně- i větněčlenské, přesto byla každá varianta testovací položky<sup>65</sup> v rámci jedné verze dotazníku použita ke zkoumání pouze jedné proměnné, aby nedocházelo ke smíšení kritérií a metoda zůstala transparentní ve smyslu: 1 varianta testovací položky – 1 proměnná (např. aktuální členění) – 1 nastavení (např. fokus).<sup>66</sup> Např. u kontextu *Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.* je jedna z variant realizována takto: *Dvacet nových zaměstnanců přijali minulý měsíc do naší firmy personalisté. Jejich pracovitost je nepřekonatelná.* Umístění této konkrétní varianty testovací položky je spojeno výhradně se zkoumanou proměnnou aktuálního členění a fokusovým nastavením. V případě jiné varianty tohoto kontextu je nastavení jiné: varianta *Personalisty bylo minulý měsíc do naší firmy přijato dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.* je zaměřena na zkoumání větněčlenské platnosti, a to konkrétně v předmětovém nastavení.

V případě, že by respondentů nebyl dostatek, bylo samozřejmě možné uvažovat o tom, zda na úkor popsané transparentnosti vytěžit data i z hlediska druhé proměnné, než která se v rámci jednotlivých variant uplatňovala, nicméně taková situace nenastala, tudíž transparentnost dat zůstala zachována.

---

zřejmé, že i další větné členy mohou být ukotveny ve valenčním rámci slovesa, např. *Petr bydlí daleko/v bytě.*

<sup>65</sup> K rozdílu mezi testovací položkou a variantou testovací položky, jakož i k samotnému designu testování viz dále.

<sup>66</sup> Nastavení varianty kontextu je předmětem následující kapitoly, proto ho zde uvádíme jen pro úplnost a podrobněji se jím na tomto místě nezabýváme.

#### 4.2.4 Testovací kontexty

Jak bylo řečeno již výše, testovací položky byly uměle vytvořeny tak, aby umožňovaly testovat navazování vztahů v textu. Při testování nebylo možné využít kontexty autentické, neboť současné především publicistické texty<sup>67</sup>, které přicházely v úvahu, nikdy nejsou nejednoznačné, neboť je zjednodušuje společný zkušenostní komplex, zjevný rámec situace, k níž odkazují. Druhým důvodem je, že ke každému testovacímu kontextu bylo nutné kvůli dotazníkovému šetření sestavit čtyři, resp. tři varianty, což s autentickými texty nebylo možné provést tak, aby nebylo patrné, že se jedná o kontexty neautentické, dotvořené. Nejednoznačné kontexty byly z těchto důvodů vytvořeny uměle. Schematicky bychom mohli strukturu každé testovací položky zaznamenat takto  $[N1_{i1}, N2_{i2}, X]_{V1} [(P_i), X]_{V2}$  a definovat ji pomocí několika charakteristik.

Každá testovací položka je dvěma samostatnými (ve smyslu oddělenými) výpověďmi, protože čím širší kontext, tím více poskytuje informací o preferovaném navazování a to je při testování nežádoucí. Bylo tedy nutné omezit rozsah testovací položky zejména v antecedentní části. Z testování jsou vyloučeny souvětné kontexty typu  $[[V1 - V2]_{\text{souvětí}}]$ , například *Karkulka měla navštívit babičku, ale sežral ji vlk.*, ale kontexty typu  $[[V1][V2]_{\text{souvětí}}]$ , např. *Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.*, v nichž se vyskytuje souvětí jako jedna část minimálního kontextu, jsou zachovány. Kontexty, v nichž by souvětnou formu měla výpověď první, tj. typu  $[[V1]_{\text{souvětí}}[V2]]$ , nejsou v materiálu zastoupeny proto, že se snažíme utvářet kontexty co nejjednodušší, aby v nich bylo co nejméně indicií pro navazování koreferenčního řetězce (především s ohledem na sdílený zkušenostní komplex). S tím souvisí i další aspekt: aby se s dalším rozšířením první výpovědi nerozšiřoval prostor pro nekontrolované faktory, které by mohly navazování koreferenčního řetězce ovlivnit.

V první výpovědi jsou dvě jmenné fráze<sup>68</sup> (N1, N2), které poskytují možnost založení koreferenčního řetězce, jsou potenciálními antecedenty (označené jako  $i_1$  a  $i_2$ ). Vzhledem k designu výzkumu, který se zaměřuje na aktuální členění a větněčlenskou

<sup>67</sup> Nejčastěji bychom potenciálně využitelné kontexty mohli nalézt v novinových titulcích, v nichž v důsledku omezeného prostoru dochází k povrchovému vypouštění některých výrazů.

<sup>68</sup> Termín jmenná fráze používám v tomto kontextu jako označení pro plnovýznamová vyjádření jmenné povahy především v první větě. Z formálního hlediska by za jmennou frázi bylo považováno i zájmeno v druhé větě, ale v tomto textu k němu takto nereferujeme.

platnost, je nutné umístit v rámci první proměnné tyto jmenné fráze do topiku a ohniska, v rámci druhé proměnné do podmětu a předmětu. K tomu blíže v oddíle Proměnné (4.2.3).

Druhá výpověď obsahuje minimálně jedno zájmeno (P) – povrchově vyjádřené i nevyjádřené, ale hloubkově přítomné ve struktuře věty (ve schématu naznačeno kulatými závorkami). Toto zájmeno je postcedentem koreferenčního řetězce s potenciálními antecedenty v první výpovědi (N1, N2). Některé typy zájmen byly z testování vyloučeny – zejména zájmena ukazovací (k nim viz dále) nebo zvrtná, neboť tato zájmena vykazují z hlediska navazování koreferenčního řetězce zřetelné preference (viz Zájmenná koreference, 4.2.5).

Koreferenční řetězec mezi jednou ze jmenných frází v první výpovědi a zájmenem ve výpovědi druhé není zřejmý ze samotného kontextu ani ze sdíleného zkušenostního komplexu mluvčích češtiny. Minimální kontext je tedy z hlediska utváření koreferenčního řetězce potenciálně víceznačný.

Ve výpovědích mohou být i další větné členy, některé z nich dokonce podmiňují větu jako takovou (např. predikát a jeho valenční doplnění). Tyto větné členy jsou ve schématu označeny jako X. Není vyloučeno, že jedním z těchto dalších větných členů je třetí jmenná fráze. Vzhledem k tomu, že v češtině se uplatňuje volný slovosled, jsou větné členy X umísťovány do takové pozice, která není příznaková.

Každý kontext existuje ve 3 variantách rozlišených první výpovědí, nejlépe je zachycuje následující tabulka J:<sup>69</sup>

aktuální členění	větněčlenská platnost
[N1 – T, Po; N2 – F, Pt] <sub>V1</sub> [(P <sub>i</sub> )] <sub>V2</sub> nastavení topikové	[N1 – T, Po; N2 – F, Pt] <sub>V1</sub> [(P <sub>i</sub> )] <sub>V2</sub> nastavení podmětové
[N1 – F, Po; N2 – T, Pt] <sub>V1</sub> [(P <sub>i</sub> )] <sub>V2</sub> nastavení fokusové	[N1 – T, Pt; N2 – F, Po] <sub>V1</sub> [(P <sub>i</sub> )] <sub>V2</sub> nastavení předmětové

**Tabulka J: schéma testovacích kontextů**

V rámci zkoumání jednotlivých proměnných variují vždy pouze hodnoty aktuálně zkoumané proměnné. V případě aktuálního členění zůstává tedy zachováno rozložení jmenných frází v rámci pozic podmětu a předmětu, ale mění se postavení v T nebo v F. N1

<sup>69</sup> Pro přehlednost a z praktických důvodů používám zkratky T = topik, F = fokus, Po = podmět, Pt = předmět, P = zájmeno.



má v první variantě kontextu platnost podmětu a postavení v topiku, zatímco ve druhé variantě má stále platnost podmětu, ale postavení ve fokusu. V případě větněčlenské platnosti je tomu naopak, zachováno zůstává rozložení aktuálněčlenské a postavení větněčlenské se proměňuje: N1 má v první variantě kontextu stejné rozložení jako v případě první varianty v rámci aktuálního členění, tedy platnost podmětu a postavení v topiku, ve druhé variantě má stále postavení topikové, ale větněčlenská platnost se změnila na předmět. Tyto variace v rámci hodnot jedné z proměnných byly již nastíněny výše, proto zde pouze pro větší zřetelnost uvádím v následující tabulce K i čtyři varianty konkrétní podoby lexikálního obsazení<sup>70</sup>.

aktuální členění	větněčlenská platnost
$[N1_{i1} - T, Po; N2_{i2} - F, Pt]_{V1} [(P_i)]_{V2}$ Pan listonoš <sub>N1,T,Po</sub> dnes doručil objednaný časopis <sub>N2,F,Pt</sub> . Zpozdil se (on).	$[N1_{i1} - T, Po; N2_{i2} - F, Pt]_{V1} [(P_i)]_{V2}$ Pan listonoš <sub>N1,T,Po</sub> dnes doručil objednaný časopis <sub>N2,F,Pt</sub> . Zpozdil se (on).
$[N1_{i2} - F, Po; N2_{i1} - T, Pt]_{V1} [(P_i)]_{V2}$ Objednaný časopis <sub>N2,T,Pt</sub> dnes doručil pan listonoš <sub>N1,F,Po</sub> . Zpozdil se (on).	$[N1_{i2} - T, Pt; N2_{i1} - F, Po]_{V1} [(P_i)]_{V2}$ Panem listonošem <sub>N1,T,Pt</sub> byl dnes doručen objednaný časopis <sub>N2,F,Po</sub> . Zpozdil se (on).

**Tabulka K: testovací kontext s notací**

Následující příklady (a–g) ukazují typy kontextů, které tvoří položky dotazníku. Pro ilustraci jsou uváděny i některé kontexty, které nejsou v dotaznících hodnoceny jako testovací. Zájmena ve druhé výpovědi jsou buď osobní (příklady a–d, g) nebo

<sup>70</sup> Na tomto místě se zdá být užitečné připomenout distinkci mezi termíny testovací položka a varianta testovací položky. Jako testovací položku označujeme souhrn všech variant jednoho lexikálního obsazení, tj. všechny tři (čtyři) varianty. Jednotlivou variantou je konkrétní realizace lexikálního obsazení v rámci hodnot dané proměnné. Testovací položka tedy označuje jednotku abstraktní, a proto, když k ní referujeme, užíváme vždy variantu T–Po. V případě konkrétní testovací položky v Tabulce B jsou jednotlivými variantami tato 3 vyjádření: *Pan listonoš dnes doručil objednaný časopis. Zpozdil se.*; *Objednaný časopis dnes doručil pan listonoš. Zpozdil se.*; *Panem listonošem byl dnes doručen objednaný časopis. Zpozdil se.* K celému lexikálnímu obsazení referujeme souhrnně jako k testovací položce *Pan listonoš dnes doručil objednaný časopis. Zpozdil se.*

přivlastňovací (*e, f*), v přímých (*a, b, c, d, e, f*) i nepřímých pádech (*b, g82*), ve všech jmenných rodech a číslech, povrchově vyjádřená (*b, d–g*) i nevyjádřená (*a–d*). Některé morfologické charakteristiky jsou ovšem zastoupeny jen v jednom z potenciálních koreferenčních řetězců (např. neutrum plurálu – příklad *c*). V některých případech není jednoznačné, jaký typ zájmena je elidován – např. osobní vs. ukazovací zájmeno. V takových případech dáváme přednost interpretaci zájmena osobního, protože zájmeno ukazovací bývá z hlediska preference utváření koreferenčního řetězce tendenční.<sup>71</sup>

- a. *Galileo veřejně podporoval Koperníka. Věřil (on), že Země se otáčí kolem Slunce.*
- b. *Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila (ona) schůzku.*
- c. *Petra musela oloupat jablka. Byla (ona) otrávená.*
- d. *Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil (on) ho.*
- e. *Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.*
- f. *Pittsburští Tučňáci odrazili newyorské Rangery. Jejich fanoušci byli překvapeni.*
- g. *Národní knihovna darovala knihy fakultě. Bylo to pro ni výhodné.*

V případě, že v jednom kontextu jsou ve druhé větě dva zájmenné výrazy, jako je tomu například ve výpovědi *d*, je otázka k tomuto kontextu v dotazníku sestavena tak, aby bylo zřejmé, který ze dvou potenciálních koreferenčních řetězců je předmětem dotazu: *Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil (on) ho.* > *Kdo/Co proslavil koho/co? > moderní dům architekta Gaudího//architekt Gaudí moderní dům.*

#### 4.2.5 Zájmenná koreference

Jak již bylo zmíněno, zájmennou koreferenci zkoumáme v dotazníkovém šetření v případech povrchově vyjádřeného i nevyjádřeného zájmena. Nerozlišujeme, zda je toto zájmeno povrchově vyjádřené obligatorně nebo potenciálně, protože to nezávisí ve většině

<sup>71</sup> Tendencnost chápeme jako preferenci jedné z možností volby antecedentu, konkrétně v tomto případě se týká preference ukazovacího zájmena navazovat koreferenční řetězec se jmennými frázemi určitého postavení (viz oddíl Ludmila Zimová, 3.3).

případů na zájmenu samotném, ale na valenci slovesného predikátu. Nicméně povrchově vyjádřených i povrchově nevyjádřených zájmen zařazených do dotazníku se týkají určitá omezení.

Z testování byly z různých důvodů vyloučeny následující skupiny zájmen:

1. Zájmena ukazovací: podle Zimové (1994: 94) „užije autor jazykového projevu [tvar ukazovacího zájmena *ten*] tehdy, je-li v antecedentní větě více pojmenování, k nimž by bylo možno zájmeno osobní vztáhnout, a odkazuje-li k pojmenování poslednímu“. Příklad, který Zimová uvádí, se nám zdá velice výstižný, proto ho uvádíme i my: *Z pěti obyvatel hořícího domu se podařilo tři snést po žebřících do bezpečí. Ve vyšších patrech zůstal [údržbář J. K.]<sub>i1</sub> a [nezletilý P. Š.]<sub>i2</sub> [Toho]<sub>i2</sub> se po marném úsilí nepodařilo zachránit.* (ibid: 94).
2. Zájmena vztažná/tázací: vzhledem k tomu, že v dotazníku je u každé testovací položky tázací zájmeno užito v otázce (viz oddíl Dotazník, 4.2.9.1), nebylo by vhodné, aby tento typ zájmen byl součástí otázek i testovacích kontextů samotných.<sup>72</sup>
3. Zájmena zvrtná: reference zvrtných zájmen je ex definitione stanovena ve prospěch podmětu. Zvrtné zájmeno lze podle českých pravidel použít pouze tehdy, když vyjadřujeme „identitu původce děje a objektu tímto dějem zasahovaného“ (MČ 2 1986: 90), např. *Půjčil jsem si pěknou knížku.* V případě přivlastňovací varianty zvrtného zájmena platí podobná podmínka: „posesivní reflexivum (adjektivní zájmeno *svůj*) vyjadřuje identitu agentu/nositele děje a vlastníka objektu tímto dějem zasahovaného“ (ibid: 91), např. *Uklid' si své věci.*
4. Zájmena neurčitá: ze své podstaty odkazují „neurčitě“, například ve struktuře *Bylo tam velké množství tanečníků, ale nikdo (z nich) se mnou nechtěl tancovat.* je jasné, že zájmeno *nikdo* odkazuje ke skupině tanečníků známé z předchozí věty, má tedy „konkrétního, ale nevybraného“ (Nědolužko 2011: 99) referenta. Tento typ vztahu považujeme za euforické odkazování. Podobnými případy se

<sup>72</sup>

Zájmena ukazovací a vztažná/tázací vykazují zřetelné preference v navazování koreferenčních řetězců, které by mohly ovlivnit respondentovy odpovědi.

podrobně zabývá Anna Nědolužko při rozlišování specifického a generického typu textové koreference (2011: 98n.).

5. Zájmena záporná: distinkci mezi zájmeny kladnými a zápornými považujeme za samostatné kritérium hodnocení zájmen nezávislé na zařazení do skupin osobní, ukazovací atd. Proto bychom i zájmena typu *nikdo*, *nic* považovali za zájmena osobní (i když s negativní sémantikou) a vyloučena z dotazníkového šetření by nebyla. Nicméně ze své podstaty odkazují záporná zájmena k bezprvkové množině možných referentů a jejich koreferenční řetězec je prázdný, tudíž není možné se jím zabývat v takovém dotazníkovém šetření, jako je to naše.

Zájmena ve druhé větě minimálního kontextu jsou tedy v dotazníkovém šetření omezena na osobní a přivlastňovací. Kontexty byly sestavovány tak, aby byly zachyceny všechny možnosti přítomnosti osobních a přivlastňovacích zájmen ve výpovědích. Je zřejmé, že většina osobních zájmen je nevyjádřená, protože je představují nevyjádřené podměty, všechna zájmena přivlastňovací jsou naopak povrchově vyjádřena, neboť není možné je povrchově nevyjádřit. Zájmen osobních je větší množství, neboť výpověď vždy obsahuje podmět (zá)jmenné povahy. Ze stejného důvodu je v některých výpovědích zájmen více.<sup>73</sup> Následující tabulka L poskytuje přehled sledovaných dotazníkových položek včetně informace o tom, kolik zájmen a jakých typů a forem vyjádření obsahují.

	kontext	počet zájmen	typ zájmena	povrchově vyjádřené
1	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.	2	osobní osobní	N A
2	Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.	1	osobní	N
3	Pan listonoš dnes doručil objednaný časopis. Zpozdil se.	1	osobní	N
4	Liberální voliči podporují demokratické kandidáty. Bojují za občanské svobody, volný trh a zmírnění chudoby.	1	osobní	N
5	Palice odrazily útok seker. Jejich zástupci bojovali statečně.	1	přivlastňovací	A

<sup>73</sup> Do počtu zájmen ve výpovědi jsou započítána jen ta, která jsou předmětem zkoumání navazování koreferenčního řetězce a v dotazníku je požadováno explicitní vyjádření toho, k jakému antecedentu přísluší (tj. jsou součástí otázky a antecedent se objevuje v možnostech odpovědi).

6	Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země.	1	přivlastňovací	A
7	Paní doktorka netrpělivě očekává snaživou kolegyni. Její pacient není připraven na operaci.	1	přivlastňovací	A
8	Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.	1	osobní	N
9	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.	2	osobní osobní	N A
10	Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.	1	přivlastňovací	A
11	Petr řídí Volkswagen. Jezdí rychle.	1	osobní	N
12	Pittsburští Tučňáci předstihli newyorské Rangery. Jejich fanoušci byli překvapeni.	1	přivlastňovací	A
13	Policie vyšetřuje korupci. Její vliv ve veřejném prostoru je nezpochybnitelný.	1	přivlastňovací	A
14	Pořadatel nejprve uvítal na konferenci zahraničního profesora. Měl radost, že se potkali.	1	osobní	N
15	Národní knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	1	osobní	A
16	Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.	1	přivlastňovací	A
17	Fotbalisté Bohemians pozvali na zápas kluky. Holky je obdivovaly.	1	osobní	A
18	Začínající autor na veletrhu představuje svůj první román. Velký zájem o něj nastartovala televizní reklama.	1	osobní	A
19	Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil ho.	2	osobní osobní	N A
20	Tomáš pozval na zápas Michala. Byl velký fanoušek fotbalu.	1	osobní	N

**Tabulka L: sledované položky a zájmena<sup>74</sup>**

Zároveň jsou do testování zapojeny jen ty kontexty s povrchově vyjádřeným zájmenem, v nichž je povrchové vyjádření zájmena neutrální variantou a neslouží ke zdůraznění nebo subjektivní hierarchizaci větné propozice, aby byly tyto potenciální vlivy z dotazníkového šetření eliminovány. Z tohoto omezení vyplývá, že v položkách dotazníku nenacházíme kontexty, v nichž by podmět byl vyjádřen povrchově přítomným zájmenem

<sup>74</sup> Položky v rámci vlastního dotazníkového výzkumu jsou označovány číslicemi, zatímco v případě testování přijatelnosti byly označovány písmeny. Tato nekonzistence je motivována faktem, že se z části nejedná o stejné položky, neboť některé kontexty, které byly součástí testování přijatelnosti, jsou například upravovány (viz oddíl Testování přijatelnosti, 4.2.7).

(jako je tomu v příkladech *i*, *j*, *k*), protože taková struktura není v češtině neutrální a zároveň takový kontext už není z hlediska navazování koreferenčních řetězců nejednoznačný. Zároveň se zdá, že případy, kdy osobní zájmeno ve funkci podmětu je povrchově vyjádřené a jeho vyjádřením se zjednotí koreferenční řetězec, jsou přirozenější v případech, kdy existuje odporovací vztah mezi oběma větami (příklad *j*).<sup>75</sup>

- h. *[Karolína]<sub>i1</sub> chtěla osobně navštívit [babičku]<sub>i2</sub>. [Pro]<sub>i1/i2</sub> odjela do Prahy.*
- i. *?[Karolína]<sub>i1</sub> chtěla osobně navštívit [babičku]<sub>i2</sub>. [Ona]<sub>i2</sub> odjela do Prahy.*<sup>76</sup>
- j. *[Karolína]<sub>i1</sub> chtěla osobně navštívit [babičku]<sub>i2</sub>. [Ona]<sub>i2</sub> ale odjela do Prahy.*
- k. *[Karolína]<sub>i1</sub> chtěla osobně navštívit [babičku]<sub>i2</sub>. [Ta]<sub>i2</sub> ale odjela do Prahy.*

Zdá se, že v kontextech typu *j* osobní zájmena vykazují preferenci k navazování koreferenčního řetězce stejnou jako zájmena ukazovací – srovnajme s příkladem *k*. Podobný rys ukazuje i Zimová (1994: 76), když srovnává elizi podmětu a jeho vyjádření osobním zájmenem na příkladech *Karel dává za pravdu Janovi, neboť on chce všechny smířit.* a *Karel dává za pravdu Janovi, neboť chce všechny smířit.* Podle Zimové je možné v prvním kontextu za antecedent považovat obě jmenné fráze (*Karel*, *Jan*) rovnocenně, zatímco v druhém kontextu s elidovaným podmětem odkazuje jen k jedné z nich, *Karlovi*, a *Jana* jako antecedent vůbec nepřipouští.

Toto pojetí není podle našeho názoru konzistentní s výkladem Ludmily Zimové, který se týká obdobných příkladů uváděných v jejím textu (1994: 83), ani dalších jejích příkladů, které jsou v této práci uvedeny jako příklady *l–n*. Ve výkladu k nim Zimová tvrdí, že v případě elidovaného podmětu není reference jednoznačná.

Toto pojetí není adekvátní ani vzhledem k jazykové realitě, protože není možné jednoznačně stanovit referenci jazykově formálně nejednoznačných anaforů bez provedení takového výzkumu, který jsme provedli my. Tudíž Zimová nemá dostatečné podklady k tomu, aby mohla tvrdit, že v kontextu s elidovaným podmětem je koreferenčně

<sup>75</sup> Tvzení týkající se hodnocení utváření koreferenčního řetězce v případech povrchově vyjádřeného zájmeného podmětu nejsou výsledkem žádného testování, pouze introspekce.

<sup>76</sup> Vzhledem k tomu, že kontext *j* bychom považovali za negramatický, nejsme schopni posoudit, nakolik je koreferenční řetězec jednoznačný. Domníváme se ale, že možnost odkazování na *[Karolína]<sub>i1</sub>* je velice oslabena, pokud ne zcela vyloučena.

specifický antecedent, ani k tomu, aby tvrdila, že existuje tendence navazovat koreferenční řetězec tendenčně ve prospěch jedné ze jmenných frází.

Z toho důvodu nemůžeme ani souhlasit s tvrzením, že „[j]e-li v antecedentní větě více substantivních pojmenování téhož gramatického rodu a čísla, která by při pronominalizaci, resp. elizi podmětu v postcedentní větě přicházela v úvahu jako antecedent a která by i z hlediska lexikálněsémantické spojitelnosti slov mohla plnit funkci podmětu postcedentní věty, užije se k realizaci postcedentu plnovýznamového pojmenování“ (Zimová 1994: 165) především proto, že potenciální nejednoznačnost nemusí být nejednoznačností z hlediska mluvčího, ani příjemce. Kontexty, v nichž v popsané situaci dochází k užití zájmena, jsou naprosto běžné.

#### 4.2.6 *Nejednoznačné navazování*

Nejednoznačnost navazování koreferenčních řetězců je v autentické komunikaci psané i mluvené nepřírozená, neboť existuje několik faktorů, které i potenciálně nejednoznačně navázaný koreferenční řetězec zjednoduší, např. společný zkušenostní komplex (znalost situace), širší textový kontext.

Je samozřejmé, že výsledky dotazníkového výzkumu mohou být zkresleny v důsledku toho, že testují nejednoznačně navazované koreferenční řetězce, tedy jazykovou situaci, která se v autentické komunikaci nevyskytuje. Nicméně se tento způsob testování jevil jako nejpraktičtější.

Princip využití nejednoznačných kontextů umožnil testovat, jaké primární čtení respondent vykazuje, tedy který koreferenční řetězec je pro něj přirozený, pokud nemá žádná další vodítka než svou metajazykovou zkušenost, tj. pokud mu chybí indicie založené především:

- a. na mimojazykové znalosti a zkušenosti respondentů: například pro interpretaci koreference v kontextu *Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.* může být klíčové, zda v rámci mimojazykové znalosti a zkušenosti bývají pracovníci spíše nově přijatí zaměstnanci, nebo personalisté,
- b. na gramatických vlastnostech struktury, které by mohly signalizovat koreferenční řetězec, kdyby se nejednalo v tomto směru o nejednoznačný kontext, srovnajme například kongruenci tvarů osobních zájmen v příkladech 1–n (příklady

a interpretace koreferenčních řetězců podle Zimová 1994: 71). Kromě toho existují ještě další faktory potenciálně zužující (ovlivňující) interpretaci koreferenčního řetězce, jako jsou např. ukazovací zájmena, o nichž jsme se zmínili výše.

- l. *Hana šla převzít službu od [primáře Vojtěcha]<sub>i</sub>. V kanceláři však zjistila, že on<sub>i</sub> bude mít službu až zítra.*
- m. *Hana<sub>i</sub> šla převzít službu od primáře Vojtěcha. V kanceláři však zjistila, že ona<sub>i</sub> bude mít službu až zítra.*
- n. *Hana<sub>i1</sub> šla převzít službu od [primáře Vojtěcha]<sub>i2</sub>. V kanceláři však zjistila, že [PRO]<sub>i1/i2</sub> bude mít službu až zítra.*

Všechny tyto potenciální vlivy byly odstraněny a testovací kontext je tedy nejednoznačný.



#### 4.2.7 Testování přijatelnosti

Vzhledem k tomu, že položky byly vytvořeny uměle, bylo před zahájením vlastního výzkumu navazování koreferenčních řetězců nutné provést testování přijatelnosti těchto umělých kontextů pro rodilé mluvčí češtiny.

Dvaceti respondentům byl předložen dotazník, v němž hodnotili všechny varianty testovacích položek na sedmibodové škále od 1 = „naprosto přijatelné, normální české vyjádření“ do 7 = „není normální, přijatelné české vyjádření“. Instrukce, které respondenti před vyplněním dotazníku obdrželi, uvádíme v plném znění:

*Vážení respondenti,*

*tento dotazník slouží jako podklad mé magisterské práce a jeho předmětem je hodnocení přijatelnosti, správnosti českých vyjádření.*

*Na sedmibodové škále zatrhněte vždy pouze jeden ze stupňů od 1 (naprosto přijatelné, normální české vyjádření) do 7 (není normální, přijatelné české vyjádření) podle Vašeho mínění. Pokud si dovedete představit situaci, v níž by dané vyjádření bylo naprosto přijatelné, znamená to označit 1, pokud podle vašeho názoru žádná taková situace neexistuje, znamená to 7.*

*Nehodnoťte, prosím, jestli je jasné, co vyjádření znamená. Může se na první pohled zdát, že některá vyjádření se opakují, ale není tomu tak. Snažte se odpovídat co nejrychleji, své odpovědi neměnit a nevracet se k již zodpovězeným položkám "pro inspiraci a srovnání".*

*Vyplnění dotazníku je anonymní, nezabere více než 15 až 20 minut a uvedené údaje slouží pouze výzkumným účelům a nebudou předány žádné třetí straně. V případě dotazů nebo zájmu o výsledky mě neváhejte kontaktovat na [alena.poncarova@gmail.com](mailto:alena.poncarova@gmail.com).*

*Velice Vám děkuji za ochotu poskytnout odpovědi. Alena Poncarová*

Testování přijatelnosti kontextů přináší několik úskalí, která většinou vyplývají z nejzávažnějšího z nich, a tím je problematika stanovení kritéria/kritérií přijatelnosti. Kritériem má podle zadání být normálnost, přijatelnost českého vyjádření. Nicméně respondentům neposkytujeme žádný praktický návod, jak tuto normálnost, přijatelnost vs. nenormálnost, nepřijatelnost určit, kromě nabádání k představování si komunikačních

situací, v nichž by dané vyjádření mohlo být použito. Z toho plyne riziko, že respondenti nemají stejně nastavenou hranici mezi normálností, přijatelností a nenormálností, nepřijatelností.

Dalším úskalím je to, že normálnost, přijatelnost bude ztotožňována s podobnými pojmy, které se v takovémto typu výzkumu obvykle vyskytují, jako jsou např. gramatičnost, adekvátnost, neutrálnost apod. Vzhledem k tomu, že objektivní kritéria pro určení a obecně platné definování normálnosti, přijatelnosti, gramatičnosti, ... neexistují, toto riziko je tedy stále aktuální a při interpretaci výsledků ho musíme mít na zřeteli.

Jednotlivé varianty testovacích položek byly v dotazníku uspořádány do tří skupin tak, že v první skupině byly varianty topikové (T) a podmětné (Po), ve druhé varianty fokusové (F) a ve třetí varianty předmětné (Pt). Varianty T a Po splývají, jedná se v rámci jednoho lexikálního obsazení o stejné struktury (viz tabulka K).

#### 4.2.7.1 Výsledky testování přijatelnosti

Celkem bylo testováno 16 lexikálních obsazení, každé z nich ve třech různých variantách s ohledem na hodnoty proměnných (aktuální členění, větný člen). Těchto 48 testovacích položek hodnotilo celkem 20 respondentů. Kompletní výsledky jsou zachyceny v příloze Výsledky testování přijatelnosti (10.3). Každý kontext je kódován písmenem A–P a číslice 1 až 3 označuje variantu konkrétního lexikálního obsazení, jehož se výsledky týkají. Kompletní přehled testovaných kontextů uvádí příloha Kontexty testování přijatelnosti (10.2), v případě ukázkového kontextu A vypadá kódování takto:

kód	varianta lexikálního obsazení	varianta proměnné
A1	Knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	T, Po
A2	Fakultě darovala knihy knihovna. Bylo to pro ni výhodné.	F
A3	Knihovnou byly knihy darovány fakultě. Bylo to pro ni výhodné.	Pt

**Tabulka M: kódování kontextů (přijatelnost)**

Výsledkem testování přijatelnosti je přehled kontextů řazený od těch, které jsou hodnoceny pozitivně, až po ty, které v testování přijatelnosti získaly negativní hodnocení.

Data byla podrobena zkoumání z hlediska středních hodnot průměru a mediánu, a také statistickému testování pomocí Studentova t-testu, kterému se budeme věnovat později. Na základě hodnot průměru a mediánu byly varianty lexikálních obsazení

uspořádány do několika skupin na základě následujících kritérií: a) hodnota průměru, b) hodnota mediánu, c) hodnocení v rámci trojice variant jednoho lexikálního obsazení. Na základě těchto kritérií byly jednotlivé skupiny vymezeny takto:

typ	kritérium	počet variant	varianty	počet kontextů	kontext	akce
I.	průměr i medián do 4.00	38	A1–D1, F1–M1, O1, P1; A2–P2; D3, H3–N3	7	D, H, I, J, K, L, M	zachovány jako testovací
II.	průměr i medián 4.00–4.50	3	B3, G3, P3	3	B, G, P	zachovány jako testovací
III.	průměr nebo medián 4.50–5.00, jediné z trojice	3	N1; F3, O3	3	F, N, O	zachovány jako testovací, upraveny po vzoru lépe hodnocených
IV.	více variant v trojici hodnoceno negativně (nad 4.50)	2	E1; E3	1	E	přesunuty do výplňkových kontextů, nahrazeny novým testovacím kontextem
V.	průměr i medián negativní (nad 5.00)	2	A3, C3	2	A, C	přesunuty do výplňkových kontextů, nahrazeny novým testovacím kontextem

**Tabulka N: přijatelnost (akce)**

První tři sloupce tabulky N představují přehled konkrétních položek a jejich zařazení do vymezených skupin na základě jednotlivých výsledků. Poslední tři sloupce v pořadí zachycují akce, které s lexikálními obsazeními (tedy s celými trojicemi variant) byly provedeny. Je zřejmé, že v prvním typu (průměr do 4.00) jsou jako testovací zachovány jen ty kontexty, u nichž nenastal ani u jedné ze tří variant případ typu II–V, neboť v rámci vlastního testování chceme zachovat jednotu lexikálního obsazení testovacích položek tak, aby byl eliminován nežádoucí vliv případných lexikálních odlišností v rámci jednoho zkoumaného kontextu.

Hranice 4.00 u typu I byla zvolena na základě přesného středu sedmibodové hodnotící škály. Do tohoto typu patří varianty A2–P2, tedy všechny varianty fokusové. Tento fakt vyplývá pravděpodobně z toho, že fokusová struktura zachovává neutrální

vyjádření agentu v podmětu, ale čeština je typická svým volným slovosledem, takže umístění tohoto Ag/Po ve fokusovém postavení nezpůsobí hodnocení kontextu jako nepřijatelného.

U typu II (kontexty *B, G, P*) byla hranice 4.50 stanovena arbitrárně, ale v rámci únosné tolerance od středové hodnoty. V rámci sedmibodové škály představuje směrodatná odchylka hodnotu vyšší než 2.0, přesně 2.160. Z toho vyplývá, že 95 % přiřazených hodnot se (za předpokladu normálního Gaussova rozdělení) vyskytuje v rozmezí 1.84–6.16. V rámci tohoto rozmezí se zdá hodnota 4.5 být adekvátní a nezkreslující, takže kontexty typu II byly pro testování navazování koreferenčních řetězců zachovány v původním lexikálním obsazení. Typy I a II se zdají být velice podobné, nicméně rozdělovat je má smysl, neboť vykazují různé hodnoty v měření přijatelnosti, a také aby bylo zřejmé, které kontexty jsou v rámci testování přijatelnosti bezproblémové (typ I) a které nejsou bezproblémové, ale spadají do sféry tolerance (typ II).

Podobným způsobem jsme postupovali i při stanovování hranice 5.00 v rámci typu III, který se týká kontextů, které mají v rozmezí 4.5–5.0 hodnocenu jen jednu z variant. Zde jsme ale již přikročili k úpravám kontextů, které spadají do tohoto typu, podle vzoru kontextů hodnocených pozitivně. Tyto úpravy se týkaly následujících kontextů ve všech třech jejich variantách:

kontext	původní kontext	upravený kontext
F	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak jí koupila květiny.	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.
N	Palice odrazily útok seker. Jejich majitelé byli překvapeni.	Palice odrazily útok seker. Jejich zástupci bojovali statečně.
O	Paní doktorka netrpělivě očekává vrchní sestru. Její dokumenty nejsou připraveny.	Paní doktorka netrpělivě očekává snaživou kolegyni. Její pacient není připraven na operaci.

**Tabulka O: upravené kontexty**

Prováděné úpravy vycházejí ze struktury kontextů hodnocených jako typ I, případně II, neomezují se pouze na antecedentní větu a jejich cílem je ustavení takového kontextu, který bude neutrální ve smyslu stylistickém i ve smyslu referenčním, tj. nebude podle našeho názoru existovat preference navazování na jednu z možných antecedentních frází. V případě kontextu *F* je úprava prováděna v rámci anaforické, tj. druhé, věty.

Kontext typu *Pak jí koupila květiny*. implikuje, že květiny koupila ta osoba, která byla vděčná a chtěla vyjádřit svou vděčnost touto formou, tedy existuje výrazná preference vybrat jako antecedent jmennou frázi *dlouholetá kolegyně*. V případě nahrazení druhé věty kontextu za *Pak si s ní domluvila schůzku*. taková preference nenastává, neboť není explicitně znám účel domluvené schůzky, tudíž ani volba antecedentu nemá preferovanou variantu. U kontextu *N jsme upravovali*, stejně jako v případě kontextu *F*, větu anaforickou, její nová podoba *Jejich zástupci bojovali statečně*. je podle našeho názoru méně příznaková z hlediska toho, zda je překvapen spíše vítěz, nebo spíše poražený. U kontextu *O* byla změněna kromě anaforické věty i jedna ze jmenných frází ve větě antecedentní. Vztah mezi jmennými frázemi v původní větě (*paní primárka a vrchní sestra*) může implikovat hierarchický vztah mezi nadřízenou a podřízenou pracovníci a tím pádem ovlivňovat volbu antecedentu. Se změnou jedné jmenné fráze ve větě antecedentní pak došlo ke změně i celé věty anaforické. Zastáváme názor, že při testování přijatelnosti těchto kontextů po provedení úprav by skóre, kterého dosáhnou, bylo typu I, nejhůře II, protože přesně kopírují strukturu kontextů hodnocených pozitivně.

Typy IV a V představují takové kontexty, které byly hodnoceny negativně nad 4.5 ve více než jedné variantě (typ IV) a nad 5.0 bez ohledu na počet variant kontextu (typ V). Tyto kontexty (A, C, E) byly pro vlastní testování nahrazeny novými kontexty:

kontext	původní kontext	nový kontext
A	Knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	Pan listonoš dnes doručil objednaný časopis. Zpozdil se.
C	Tomáš pozval na zápas Michala. Byl velký fanoušek.	Fotbalisté Bohemians pozvali na zápas kluky. Holky je obdivovaly.
E	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.	Pittsburští Tučňáci předstihli newyorské Rangery. Jejich fanoušci byli překvapeni.

Tabulka P: nové kontexty

Kontexty A, C a E na základě testování přijatelnosti musely být nahrazeny kontexty novými. V případě kontextu A si nejsme jistí, co na něm způsobilo negativní hodnocení, protože se nám zdál být prototypickým případem kontextu hodnoceného pozitivně až průměrně. Mohl to způsobit například fakt, že *knihovna* i *fakulta* mohou mít referenci k instituci i k budově, fakt, že ve větě je kromě těchto dvou jmenných frází i třetí jmenná fráze, která má stejný slootovorný základ jako jedna ze zkoumaných jmenných frází

(*knihy*), nebo fakt, že *knihovna* i *fakulta* představují holé větné členy, nejsou nijak rozvity. Vzhledem k této nejistotě o původu negativního hodnocení není úplně zřejmé, jakým kontextem by měl být nahrazen. Proto jsme vytvořili kontext zcela odlišný, kde jednu ze jmenných frází zastupuje člověk (*pan listonoš*) a druhou věcný denotát (*objednaný časopis*), navíc třetí jmenná fráze v této nové variantě není. Kromě toho byl původní kontext A, přestože již není testovacím kontextem, také upraven do podoby *Národní knihovna darovala...*, aby bylo zřejmé, že neexistuje přímá vazba mezi *fakultou* a *knihovnou*, a zároveň kontext působil přirozeněji.

U kontextu C jsme zachovali slovesný predikát i tematickou oblast, nicméně jsme nahradili *Tomáše* a *Michala* jmennými frázemi s obecnější denotací. Kontext E je dalším případem, u něhož si neumíme vysvětlit příčiny toho, proč je spíše negativně hodnocen. Z toho důvodu k němu přistupujeme stejně, jako jsme postupovali v případě kontextu A, tedy vytvořením naprosto odlišného kontextu. V tomto případě dochází k vyrovnaní hierarchických vztahů mezi jmennými frázemi (*pittsburští Tučňáci* a *newyorští Rangeři*) z původně nerovného vztahu mezi jmennými frázemi způsobeného pravděpodobně disproporcí mezi apelativy a proprii (*pan domácí* a *Karel*).

Vzhledem k tomu, že předpokládáme normální rozložení dat na škále 1–7 a pracujeme s řadovou proměnnou, je t-test vhodným nástrojem pro testování statistické spolehlivosti, signifikantnosti dat. Kompletní přehled výsledků t-testu je uveden v příloze T-test (10.5). Z této tabulky zachycující hodnoty pravděpodobnosti (p-values), tedy hladiny spolehlivosti naměřeného rozdělení, je patrné, že na hranici pětiprocentní hranice spolehlivosti se pohybuje jen šest kontextů, ale pouze v hodnocení přijatelnosti v rámci srovnání variant z hlediska větněčlenského. Jedná se o kontexty A, B, C, G, O a P. V případě těchto kontextů tedy můžeme předpokládat, že rozložení jednotlivých odpovědí (tedy hodnocení jednotlivých respondentů týkajících se daných kontextů) v rámci srovnání variant 1 (Po) a 3 (Pt) je signifikantní a při opakování měření bychom na úrovni pětiprocentní pravděpodobnosti získali stejně rozložená hodnocení.

Tento test musíme zkombinovat se zařazením kontextu do typů (I–V), protože jinak nevypovídá nic o tom, jak byly kontexty hodnoceny, tedy jestli varianty dostaly pozitivní, nebo negativní hodnocení. Při srovnání zjistíme, že v případě kontextů A a C, o kterých již

byla řeč výše v rámci nahrazovaných kontextů, je rozdíl mezi variantami 1 a variantami 3 diametrální, varianty 1 náleží k typu I, zatímco varianty 3 k typu V, tedy každá z těchto variant zastupuje opačný pól škály přijatelnosti. Případy kontextů *B*, *G* a *P* jsou jiného typu, protože na první pohled mezi nimi nejsou viditelné tak výrazné rozdíly. Varianty *B1*, *G1* a *P1* náleží k typu I, zatímco varianty *B3*, *G3* a *P3* k typu II. Trojice těchto kontextů jsou pro vlastní testování týkající se navazování koreferenčních řetězců zachovány v nezměněné podobě. V případě kontextu *P* se varianty 1 a 3 liší na úrovni odlišnosti typů I a III, tento kontext tedy byl zachován, ale v podobě upravené (viz výše). Nicméně tato data chápeme jako informativní, protože jsou založena na odpovědích pouze dvaceti respondentů.

S plným vědomím představených rizik byly na základě výsledků tohoto testování sestaveny dotazníky pro vlastní testování. Jejich podoba je zachycena v příloze Dotazníkové položky (10.6). Pokusili jsme se alespoň v rámci možností otestovat jednotlivé varianty a dotazníky sestavit na základě výsledků těchto předběžných testů.

#### 4.2.8 Výplňkové kontexty

Výplňkové kontexty vycházejí z náhodně vybraných autentických českých textů.<sup>77</sup> Tyto kontexty ale nejsou zachovány v původní verzi, jejich forma je upravována tak, aby byly co nejpodobnější struktuře kontextů testovacích, často byla prováděna zejména parcelace souvětí do dvou vět jednoduchých. Touto parcelací se výplňkový kontext stal ve většině<sup>78</sup> případů nejednoznačným, což ovšem není překážkou, spíše naopak – čím podobnější testovacím kontextům se podaří výplňkové kontexty vytvořit, tím lépe.

Výplňkové kontexty nepodléhají při výběru tak přísným pravidlům jako kontexty testovací, protože na ně nejsou kladeny nároky na vytvoření tří různých variant, výplňkový kontext má ve všech variantách dotazníku stejnou strukturu. Z toho důvodu jsou mezi výplňkové kontexty zařazeny minimální dvojice, v nichž působí například mimojazyková realita jako silný faktor ovlivňující interpretaci koreferenčního řetězce, srovnáme kontexty

<sup>77</sup> Většinou se jedná o novinové titulky nebo části novinových článků.

<sup>78</sup> Kontexty, které zůstávají jednoznačné i po úpravách, jsou v příloze Dotazníkové položky (10.6) označeny jako VN (výplňkové, ne-nejednoznačné).

*p–t*, které jsou jednoznačné na základě gramatických pravidel (shoda v čísle, rodě), s kontextem *o*, který jednoznačný není.

- o. *TOP 09 je nyní v Poslanecké sněmovně. O její budoucnost se neobávám.*
- p. *Češi chtějí vidět Jaromíra Jágra. Jejich touha vyprodala O2 arénu.*
- q. *Norové odmítli Evropskou unii. Ted' (oni) jí dávají Nobelovu cenu míru.*
- r. *Do České filharmonie se vrací bývalý dirigent. Jeho dva koncerty vyvolaly kladné ohlasy.*
- s. *Budova nového pavilonu v botanické zahradě bude patřit různým. Ty jsou novým logem zahrady.*
- t. *České fotbalisty čeká další bitva. Snad ona dopadne dobře.*

Odpovědi na otázky týkající se těch výplňkových kontextů, které se staly nejednoznačnými, by v případě nedostatku signifikantních dat (zejména v důsledku nedostatku respondentů) mohly sloužit jako zdroj dalších informací. Tato situace by byla problematická zejména proto, že výplňkové kontexty existují jen v jedné variantě, nejčastěji ve variantě T–Po, tudíž by množství dat nekorespondovalo s množstvím dat k proměnné získaných v rámci testovacích kontextů. Tato možnost ale nenastala, tudíž výplňkové kontexty zůstaly jen položkami doplňujícími kontexty testovací.

#### 4.2.9 Testování navazování koreferenčních řetězců

Ještě než přejdeme k vlastnímu testování, popíšeme položky, které označujeme jako sledované. Mezi sledované položky patří

- a) všechny původně vytvořené testovací kontexty, které byly v testování přijatelnosti hodnoceny pozitivně (typ I a II,<sup>79</sup> celkem 10 kontextů),
- b) kontexty, které v rámci testování přijatelnosti tvoří typ III, tj. kontexty modifikované (celkem 3 kontexty),

<sup>79</sup> Vymezení jednotlivých typů viz tabulka O.



- c) kontexty nově utvořené po vzoru pozitivně hodnocených kontextů, které nahrazují původní negativně hodnocené kontexty (nové testovací položky nahrazující kontexty typu IV a V, celkem 3 kontexty),
- d) původní kontexty, které byly v rámci testování přijatelnosti hodnoceny negativně a tvoří typ IV a V (celkem 3 kontexty),
- e) jediný autentický kontext, který na formální rovině nebyl jednoznačný, a zároveň bylo možno utvořit všechny tři vyžadované varianty: Jedná se o kontext *Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země*. Je jasné, že koreferenční řetězec sice není zřejmý z gramatické stavby kontextu, ale na základě mimojazykové situace je naprosto zřejmý koreferenční vztah mezi *Milošem Zemanem* a *jeho příznivci*. Z toho důvodu jsou pouze v případě tohoto kontextu jinak utvořeny možnosti odpovědí v dotazníkovém šetření (viz oddíl Dotazník, 4.2.9.1).

Vlastní testovací kontexty jsou zastoupeny kontexty skupiny a), b) a c), celkem 16 kontextů. Testovací kontexty tedy tvoří jen část sledovaných položek. Další sledované položky tvoří kontexty, které byly původně hodnoceny negativně (skupina d), celkem 3 kontexty), a jeden kontext autentický. Každý dotazník tedy obsahuje dvacet sledovaných kontextů.

#### 4.2.9.1 Dotazník

Klíčovým faktorem, který ovlivňuje výsledky dotazníkového šetření, je podoba dotazníku. Již bylo řečeno, že v rámci zkoumání pracujeme se čtyřmi variantami dotazníku.

##### 4.2.9.1.1 Sestavení

Vlastní dotazník je sestaven na základě několika pravidel:

- v rámci dotazníkového šetření vymezujeme kromě testovacích a výplňkových položek i sledované položky,<sup>80</sup>

<sup>80</sup> Ke vztahu testovacích a sledovaných položek viz oddíl Testování navazování koreferenčních řetězců (4.2.9).

- pořadí sledovaných vs. výplňkových položek je arbitrární, pro všechny čtyři varianty stejné, mění se jen varianta sledované položky,
- rozložení a pořadí položek zachycuje příloha Dotazníkové položky (10.6; v ní jsou sledované položky zachyceny jako T a výplňkové jako V),
- každý dotazník obsahuje 5 sledovaných položek T, 5 sledovaných položek F, 5 sledovaných položek Po, 5 sledovaných položek Pt – složení dotazníku schematicky zachycuje následující tabulka, konkrétní lexikální obsazení variant testovacích položek zachycuje příloha Sledované položky (10.4). Jednotlivé kontexty rozčleněné do skupin po pěti představují konkrétní lexikální obsazení. Aby nedošlo k situaci, že by se v jedné variantě dotazníku (Q1–Q4) vyskytlo stejné lexikální obsazení v různých variantách, jsou kontexty kombinovány do variant dotazníků tak, jak ukazuje tabulka Q. Volba příslušnosti jednotlivých variant kontextů k jednotlivým variantám dotazníku je náhodná.

	Q1 fialová	Q2 modrá	Q3 zelená	Q4 žlutá
T	kontexty 1–5	kontexty 6–10	kontexty 11–15	kontexty 6–10
F	kontexty 6–10	kontexty 1–5	kontexty 16–20	kontexty 11–15
Po	kontexty 11–15	kontexty 16–20	kontexty 1–5	kontexty 16–20
Pt	kontexty 16–20	kontexty 11–15	kontexty 6–10	kontexty 1–5

Tabulka Q: Q1–4<sup>81</sup>

- výplňkové kontexty jsou stejné v rámci všech čtyř variant dotazníku, a to včetně varianty uspořádání z hlediska T–F a Po–Pt a včetně umístění v rámci dotazníku

Testování probíhalo elektronickou formou, prostřednictvím speciálně upravené aplikace. Respondentům byl předložen tento úvodní (informační a instruktážní) text:

<sup>81</sup> Kontexty jsou v rámci testování přijatelnosti označovány písmeny, ale v rámci testování navazování koreferenčních řetězců jsou označovány číslicemi, aby bylo zřejmé, že nejsou plně totožné s kontexty, které byly testovány v rámci testování přijatelnosti.

*Vážení respondenti,*

*tento dotazník slouží jako podklad mé diplomové práce. Jeho předmětem jsou zájmena a jejich fungování v současné češtině.*

*Vyplnění dotazníku nezabere více než 30 minut. Vyplnění dotazníku je anonymní, uvedené údaje slouží pouze výzkumným účelům a nebudou předány žádné třetí straně.*

*V případě dotazů nebo zájmu o výsledky mě neváhejte kontaktovat na [alena.poncarova@gmail.com](mailto:alena.poncarova@gmail.com). V případě, že máte dotazy k vyplňování dotazníku, kontaktujte mě ještě před vyplňováním.*

*Velice Vám děkuji za ochotu poskytnout Vaše odpovědi.*

*Alena Poncarová*

*V další části se seznámíte s instrukcemi k vyplnění dotazníku: Dotazník obsahuje celkem 50 otázek. Každá otázka má 3 části: text, otázku vztahující se k textu a 2 možnosti odpovědi na tuto otázku. Text i otázku si pečlivě přečtete hned napoprvé. Na otázky odpovídejte označením vždy jen jedné odpovědi z nabídky. Odpovědi na všechny otázky jsou povinné. Časový limit pro přečtení jednoho textu a zodpovězení otázky je 20 vteřin. Zbývající čas je u každé otázky zobrazen šedým pruhem nahoře. Označené odpovědi nelze měnit. Po zatržení zvolené odpovědi se zobrazí tlačítko Další, kterým pokračujete na další položku. K odeslaným textům/otázkám se nelze vrátet.*

Po tomto úvodním informačním a instruktážním textu byly v dotazníku umístěny tři položky, na nichž se respondenti seznámili s ovládáním a fungováním aplikace (včetně časového limitu dvaceti vteřin). Každá z otázek se zobrazovala na samostatné stránce, aby byly prostředí a situace co možná nejvíce podobné tomu, jaké byly v případě vlastních dotazníkových položek.

*Nyní budou následovat tři instruktážní položky. Na nich si vyzkoušíte, jak rychle musíte odpovídat na otázky a jak ovládat program.*

*Jako počáteční hranici stanovili autoři rok 1922. Tehdy začala vycházet série Josefa Lady. Co se stalo v roce 1922?*

[ ]<sup>82</sup> *Autoři stanovili počáteční hranici.*

[ ] *Začala vycházet série Josefa Lady.*

*Petr fauloval Tomáše. Nelíbilo se mu, že hrál rukou.*

*Kdo koho fauloval?*

[ ] *Petr Tomáše*

[ ] *Tomáš Petra*

*Plzni se zranil další útočník. Do příštího zápasu je výrazně oslabená.*

*Kdo je výrazně oslabený do dalšího zápasu?*

[ ] *Plzeň*

[ ] *další útočník*

Struktura těchto instruktážních položek plně odpovídá struktuře, jakou mají položky vlastního dotazníku. Až do tohoto místa jsou čtyři varianty dotazníku totožné, poté následují vlastní dotazníkové položky, které se v rámci jednotlivých verzí dotazníku liší (viz příloha Dotazníkové položky, 10.6).

#### 4.2.9.1.2 Otázky

Každá položka má v rámci dotazníku přiřazenu vlastní otázku<sup>83</sup>. Toto řešení je jediným možným, protože různá lexikální obsazení jednotlivých položek, různé valenční struktury závislé na konkrétním slovese neumožňují formulovat obecnou otázku, na kterou by respondenti odpovídali.

Příklady jednotlivých otázek uvádíme zde:

- u. *(Roger Federer odehrál exhibiční zápas s Rafaelem Nadalem. Publikum mu tleskalo ve stoje.) Komu tleskalo publikum vestoje?*

<sup>82</sup> Hranaté závorky představují zástupný znak pro zaškrtačací políčko, které slouží k výběru jedné z nabízených možností.

<sup>83</sup> Kompletní seznam otázek a nabízených odpovědí obsahuje příloha Otázky a možnosti odpovědi (10.7).

- v. *(Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.) Kdo si domluvil schůzku s kým?*
- w. *(Vize jednotné Ameriky byla překryta novým přístupem. Apeluje na třídní nevraživost.) Co apeluje na třídní nevraživost?*
- x. *(Volební výsledky mají moc zamíchat dozorčí radou. Přímou či nepřímou závisí na hlasech voličů.) Kdo/Co závisí na hlasech voličů?*
- y. *(Autosalon propaguje nový vůz. Jeho úspěšnost závisí na reklamě.) Čí úspěšnost závisí na reklamě?*
- z. *(Pan listonoš dnes doručil objednaný časopis. Zpozdil se.) Kdo/Co se zpozdil/o?*
- aa. *(Postřelenou dívku převezli do jiné nemocnice. Je lépe hlídána.) Kdo/Co je lépe hlídáný/é?*

Sestavování otázky se, jak je patrné z uvedených příkladů, řídí několika základními pravidly.

1. Otázka se v rámci všech čtyř variant dotazníku nemění, pro daný testovací i výplňkový kontext zůstává stále tatáž.
2. Jedná se o otázku doplňovací a uzavřenou, respondent musí vybírat svou odpověď ze dvou nabízených možností.
3. Struktura položené otázky nesmí respondenty nasměrovat k jedné z nabízených možností. Z toho důvodu uvádíme v otázce například zájmena odkazující k osobám i věcem (*kdo, co*). Tuto zásadu dodržujeme naprosto striktně nejen v případě tázacích zájmen, ale i v případě shody v přičestích minulých, z toho důvodu jsou některé otázky komplikovanější, než je obvyklé, zejména tam, kde se v otázce vyskytuje přičestí vyžadující dvojí různou shodu podle navázaného koreferenčního řetězce (příklady *z, aa*). V případech, kdy obě možnosti navázání koreferenčního řetězce vyžadují jen jedno z tázacích zájmen (příklad *v, w*), uvádíme v otázce jen to jedno, neboť nemá smysl zbytečně otázku komplikovat. Pokud je možná návaznost zájmenem osobním i věcným, podle smyslu nebo formální shody, uvádíme obě varianty (příklad *x*).

4. V případě přivlastňovacích zájmen užíváme tázací zájmeno *čí* i v případě, že se přivlastňuje neživotné entitě, protože v současném úzu je podle našeho názoru takovému přivlastňování častější než typ genitivní (*koho, čeho* – příklad *y*)
5. V rámci otázky musí být všechna doplnění, která jsou součástí druhé věty minimálního kontextu zopakována. Je zřejmé, že pokud sloveso vyžaduje povrchově obligatorní doplnění, není ani možné ho vynechat (příklad *v, w, x, y*).
6. Slovosled otázky musí být neutrální a nesmí implikovat žádnou citově zabarvenou situaci, která by mohla ovlivnit respondentovo rozhodování, kterou z nabízených možností zvolí. Z toho důvodu aplikujeme v otázce princip řazení Agens – VF – Pat – (X), tj. kdo dělá co (komu, jak, ...).

#### 4.2.9.1.3 Nabízené možnosti odpovědi

Jak již bylo zmíněno, každá položka zakládá vlastní otázku a vlastní lexikální obsazení odpovědi<sup>84</sup>. Nejsnadněji popíšeme strukturu dotazníkových položek na příkladu:

*Roger Federer odehrál exhibiční zápas s Rafaellem Nadalem. Publikum mu tleskalo ve stoje.*

*Komu tleskalo publikum ve stoje?*

*[ ] Rogeru Federerovi*

*[ ] Rafaelu Nadalovi*

Sestavení dvou možností, z nichž respondenti vybírají pouze jednu, podléhalo několika pravidlům:

- Odpovědi v podstatě kopírují Danešův otázkový test (1968), simulují autentickou dialogickou situaci.
- Z logicky možných čtyř variant odpovědi (*Rogeru Federerovi, Rafaelu Nadalovi, oběma, žádnému z nich*) byly nabízeny jen dvě možnosti. Tzv. *forced-choice* byla zvolena především z toho důvodu, že jsme chtěli co nejméně upozorňovat na

<sup>84</sup> Kompletní seznam otázek a nabízených odpovědí obsahuje příloha Otázky a možnosti odpovědi (10.7).

nejednoznačnost kontextu a uvedením možnosti *oba* by podle našeho názoru došlo právě ke zvýraznění nejednoznačnosti. Varianta *žádnému z nich* je jen formální, neboť kontexty vždy implikují vazbu k antecedentu v předchozí větě.

- Pořadí variant odpovědi je náhodné a v různých variantách dotazníku zachováno, neboť případný vliv tohoto pořadí je odstíněn tím, že se napříč čtyřmi verzemi dotazníku proměňuje kontext, tudíž je (podle našich hypotéz) preferovaná odpověď střídavě první a druhá.
- Odpovědi jsou u všech položek sestavovány pokud možno co nejpodobněji tak, aby nebylo patrné, který kontext je sledovaný a který pouze výplňkový.
- Odpovědi nemají formu kompletních výpovědí, ale pouze jmenných frází. Forma výpovědní by totiž jednak v dialogu nebyla přirozená, jednak by tak byla uměle prodlužována doba, za jakou lze celý text přečíst, a o to by musel být delší časový limit u každé otázky.
- U otázek, které z důvodu přesného vymezení rolí jednotlivých jmenných frází v rámci obsazení valenčních pozic predikátu vyžadují zahrnutí více jmenných frází do každé varianty odpovědi, je pořadí těchto jmenných frází stanoveno neutrálně ve smyslu slovosledném. Z toho důvodu jsou na otázku *Kdo dává Nobelovu cenu míru komu?* varianty odpovědi sestaveny tak, aby první v pořadí byla jmenná fráze tvořící odpověď na otázku *kdo* a druhá v pořadí jmenná fráze tvořící odpověď na otázku *komu* a ne naopak. Varianty odpovědi tedy jsou: *Norové Evropské unii X Evropská unie Norům*, nikoli *Evropské unii Norové X Norům Evropská unie*, neboť tento slovosled není neutrální a implikované konotace by mohly ovlivnit volbu varianty. Jedinou výjimku z tohoto pravidla tvoří kontext *Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země.*, kde jsou nabízené varianty odpovědi slovosledně příznakové: *Kdo střídá koho? – prezidenta Václava Klause Miloš Zeman X Miloše Zemana prezident Václav Klaus*. V tomto případě máme možnost otestovat pozornost respondentů, jejich pečlivost při čtení kontextů a nabízených odpovědí spíše než navazování koreferenčního řetězce, protože to je zkresleno silným vlivem mimojazykové zkušenosti. Výpovědní hodnota tohoto měření by byla větší, kdyby kontext byl zařazen až v druhé části dotazníku (je šestnáctým kontextem v pořadí), nicméně i tak je možné sledovat, že respondenti volí variantu

*prezidenta Václava Klause Miloš Zeman* výrazně častěji než variantu druhou, která mimojazykové skutečnosti neodpovídá (261:18). Těmto výsledkům se budeme blíže věnovat v části Výsledky dotazníkového výzkumu (4.2.10).

- V případě, kdy je jmenná fráze obsáhlá, například v důsledku rozvití adjektivním přívlastkem jako v případě *[Známým architektem Gaudím] byl navržen [moderní dům ve Francouzské čtvrti]. Proslavil ho.*, není nutné zachovávat v nabízené odpovědi celou jmennou frázi, neboť by se odpověď stala nepřehlednou a hrozilo by zvýšení chybovosti v důsledku nepozorného čtení. Z toho důvodu je například v tomto konkrétním případě vypouštěn shodný přívlastek *známý* a neshodný přívlastek *ve Francouzské čtvrti*. Roli hraje i časové hledisko, o kterém jsme se zmínili výše. Z těchto důvodů náleží k otázce *Kdo/Co proslavil/o koho/co?* ne odpovědi zahrnující kompletní jmenné fráze, ale pouze některé části jmenných frází: *moderní dům architekta Gaudího X architekt Gaudí moderní dům*.
- Pokud je v rámci kontextu jmenná fráze ve vazbě s předložkou, je tato předložka i součástí nabízených variant odpovědi. Například v kontextu *Začínajícím autorem je na veletrhu představován jeho první román. Velký zájem o něj nastartovala televizní reklama.* jsou na otázku *O koho/co nastartovala zájem televizní reklama?* varianty odpovědi *o začínajícího autora X o jeho nový román*. Tento kontext je zároveň jediným, v němž existuje mezi jednotlivými variantami odpovědi koreferenční vztah: *o [začínajícího autora]<sub>i</sub> X o [jeho]<sub>i</sub> nový román*, neboť by bylo velice komplikované vyjadřovat druhou variantu odpovědi jiným způsobem a vynechání posesivního zájmena by mohlo vést k zavádějící interpretaci (*o nový román – který napsal někdo jiný?*).

Otázky i nabízené varianty odpovědi jsou zobrazeny současně s každým kontextem, není tedy žádný předěl mezi zobrazením kontextu a zobrazením otázky a nabízených variant odpovědi.

#### **4.3 Výsledky dotazníkového výzkumu**

Podle popsané metodiky jsme provedli samotné testování navazování koreferenčních řetězců. Každou verzi dotazníku vyplňovalo 71 respondentů, tj. celkem se našeho dotazníkového šetření účastnilo 284 respondentů. To, kterou variantu dotazníku respondent



dostane, vybírá sama aplikace podle toho, jaká verze dotazníku má aktuálně nejmenší počet vyplnění. Tímto způsobem je zajištěna rovnoměrná distribuce variant dotazníku (a tím i jednotlivých sledovaných položek).

Souhrnně máme k dispozici 5 680 měření (k tomuto počtu viz dále). Ne všechna z nich jsou ale použita k vlastnímu statistickému ověřování, neboť v některých případech se stalo, že respondent položku nevyplnil. Přehled zachycující tyto hodnoty je uveden v následující tabulce R:

	aktuální členění	větný člen	
<i>T/F</i>	2 810	2 796	<i>Po/Pt</i>
<i>nezaznamenáno</i>	30	44	<i>nezaznamenáno</i>
<i>sum</i>	2 840	2 840	<i>sum</i>

**Tabulka R: nezaznamenané hodnoty**

V případě druhé proměnné je počet nezaznamenaných odpovědí poněkud vyšší, nicméně i tak představuje pouze 1.549<sup>85</sup> procenta všech měření týkajících se dané proměnné (v případě aktuálního členění je tento poměr 1.068), což považujeme za zanedbatelné, neboť zaznamenání odpovědi respondentů i v případech, které v aktuální situaci nemáme k dispozici, by na celkové výsledky mělo pouze zanedbatelný vliv. Proto považujeme data uvedená v řádku T/F, resp. Po/Pt za výchozí, jsou základem pro níže uvedené závěry.

Na tomto místě představíme nejprve výsledky navazování koreferenčních řetězců týkající se obou proměnných, poté se zaměříme na navazování koreferenčních řetězců z hlediska jednotlivých lexikálních obsazení. Všechny pravděpodobnostní hodnoty (p-values) vycházejí z tzv. testu dobré shody zvaného také chí-kvadrát ( $X^2$ , chi-square). Tento test byl zvolen z toho důvodu, že umožňuje srovnávat počty výskytů jedné hodnoty s počty výskytů hodnoty druhé, na rozdíl od Studentova t-testu, který je vhodné použít spíše pro proměnné řadové nebo numerické.

<sup>85</sup> Všechna data číselné povahy uvádím se zaokrouhlením na tři desetinná místa, pokud není uvedeno jinak a pokud se nejedná o případy, kdy je taková přesnost zbytečná a data představují celá čísla.

#### 4.3.1 Aktuální členění

Aktuální členění s hodnotami T a F bylo první testovanou proměnnou. V každé variantě dotazníku je zastoupena v celkem deseti sledovaných položkách, tj. 5 lexikálních obsazení v nastavení T–Po a 5 lexikálních obsazení v nastavení F–Po (k tomuto nastavení viz Testovací kontexty, 4.2.4). Souhrnně se tedy této proměnné týká 2840 položek, tj. 4 varianty dotazníku, v každé z nich 10 položek týkajících se AVČ, 71 respondentů každé varianty dotazníku. Z tohoto celkového počtu je 1420 položek v nastavení T–Po a 1420 položek v nastavení F–Po. Po vyloučení těch měření, v nichž nebyla hodnota (T vs. F) zaznamenána, pracujeme v rámci aktuálního členění s 2810 měřeními.

Pokud sečteme odpovědi všech respondentů, v nichž sledujeme proměnnou aktuální členění, získáme tyto sumy:

	T	F
sum	1 542	1 268
chi <sup>2</sup>	26.717	
p-value	2.355e-07	

**Tabulka S: T–F sum**

Z tabulky S je zřejmé, že navazování na T je signifikantně častější než na F. Rozdíl 274 ovšem v kontextu všech měření není nijak přesvědčivý. Nicméně tato prostá tabulka nezachycuje, a ani nemůže zachytit, jak se proměňuje navazování na T a F se změnou uspořádání antecedentní věty, které je odlišné v rámci každé verze dotazníku. Abychom mohli aplikovat takovouto perspektivu, musíme zachytit rozložení T a F s ohledem na varianty dotazníku, aby bylo zřejmé, které variantě kontextu odpovídá kolik a jakých měření. Tomu slouží dvě následující tabulky. V první z nich, v tabulce T, jsou zachyceny odpovědi respondentů (T vs. F) v rámci jednotlivých variant dotazníku.

	T	F	rozdíl	p-values	p < 0.05
Q1	321	385	64	0.016	A
Q2	439	255	184	2.858e-12	A
Q3	419	284	135	3.55e-07	A
Q4	363	344	19	0.475	N
	1 542	1 268	402	2.355e-07	A

**Tabulka T: T–F (Q1–4)**

Průměrný počet měření, v nichž respondenti označují za antecedent zájmena T, je 385,5, v případě F je tento průměr nižší, 317. Počty měření odpovídajících hodnotě T a hodnotě F jsou v rámci jednotlivých dotazníkových variant rozdílné, celkový rozdíl 402 měření ve prospěch varianty T je na první pohled solidním výsledkem, na devadesátipětiprocentní hladině pravděpodobnosti si můžeme být jisti, že rozložení dat není náhodné, že při opakovaném měření by výsledek vykazoval podobné tendence, tedy že navazování na T je častější, vezmeme-li v úvahu údaje ze všech variant dotazníků jako celek.

Pokud chceme blíže zkoumat navazování koreferenčních řetězců v závislosti na variantě dotazníku, zjistíme, že tendence navazování na T není u všech převažující. Dominantní je pouze u variant dotazníku Q1–Q3. Rozložení dat u varianty dotazníku Q4 není signifikantní. Zajímavé jsou rozdíly mezi jednotlivými variantami dotazníku, které mohou vycházet z konkrétních variant lexikálních obsazení, která jsou jejich součástí. V tomto směru se zdá, že bychom čtyři varianty dotazníku mohli rozdělit do dvou skupin, první z nich představují varianty dotazníku Q1 a Q4, kde nejsou rozdíly mezi hodnotami T a F nijak extrémní, druhou skupinu tvoří varianty dotazníku Q2 a Q3, které jsou v rámci rozdílů hodnot T a F přesvědčivější (k lexikálním obsazením a jejich měření viz Jednotlivé kontexty, 4.3.3).

	T > T	T > F	F > F	F > T
Q1	208	145	240	113
Q2	112	237	202	143
Q3	222	131	153	197
Q4	225	129	215	138

**Tabulka U: T–F predikce**

Z tabulky U je zřejmé, že situace není tak jednoduchá, jak se může na základě tabulky S zdát. Jednotlivé sloupce zachycují nastavení varianty kontextu a odpověď respondentů, např. sloupec  $T > T$  představuje taková měření, v nichž nastavení varianty kontextu bylo topikové a respondenti za antecedent tento topik označovali, sloupec  $T > F$  představuje situaci, kdy v topikovém uspořádání respondenti označují za antecedent fokus. V případě, že varianta kontextu je topiková (T), ve většině případů dochází k navazování na T, ale v případě dotazníku Q2 je tomu naopak. Podobně je to i v případě nastavení F, ve většině případů dochází k navazování na F, ale v případě dotazníku Q3 je preferovanou variantou T.

Zároveň v případě varianty Q3–F není rozdíl mezi hodnotami (153 vs. 197) tak velký jako v ostatních případech, přesto jsou všechna rozložení hodnot signifikantní na pětiprocentní úrovni (kompletní hodnoty v příloze nazvané T–F podle nastavení kontextu, 10.8), tedy i v případě Q3–F, která je signifikantní ve prospěch T. K jednotlivým kontextům, které by mohly rozdíly mezi dotazníky způsobovat, viz sekci Jednotlivé kontexty (4.3.3).

Naše hypotéza týkající se aktuálního členění predikuje výsledky pouze v rámci dvou ze čtyř možných variant: v případě, že v dotazníku je varianta T, odpovědi T, v případě, že v dotazníku je varianta F, odpovědi F. Nepredikuje případy, v nichž v nastavení T respondent označí F a naopak. Hypotéza ale představuje pouze ideální konstrukt, ve skutečnosti, jak je patrné, jsou nepredikované případy časté. Vysvětlením toho, že existují případy nepredikované hypotézou, může být například fakt, že volný slovosled v češtině umožňuje relativně libovolné umístění frází ve výpovědi. Přesto můžeme v datech najít signifikantní tendenci k navazování na topikovou část výpovědi.

#### 4.3.2 *Větněčlenská platnost*

V případě větněčlenské platnosti je rozložení položek v dotazníku podobné jako v případě aktuálního členění, celkem 10 položek věnovaných této proměnné, z nich 5 v uspořádání Po–T, druhých 5 v uspořádání Po–F. I této proměnné se týká celkem 2840 položek, po odečtení nezaznamenaných měření jich je celkem 2796.

V následující přehledové tabulce V uvádíme počty všech měření rozdělené podle hodnoty větněčlenské platnosti. Rozdíl mezi hodnotami Po a Pt je o polovinu větší než mezi hodnotami T a F představený výše (Po–Pt 606, T–F 402). To samo o sobě už

naznačuje rozdíly v proměnných, naznačuje to větší spolehlivost výsledků týkajících se větněčlenské platnosti.

	Po	Pt
sum	1 701	1 095
chi <sup>2</sup>	131.343	
p-value	2.2e-16	

**Tabulka V: Po–Pt sum**

K zachycení rozdílů mezi jednotlivými variantami dotazníků Q1–4 nám poslouží opět tabulka. Stejně jako v případě aktuálního členění (tabulka T) zachycuje následující tabulka W sumy odpovědí respondentů ve prospěch podmětu (Po) a předmětu (Pt) v závislosti na verzi dotazníku.

	Po	Pt	rozdíl	p-values	p < 0.05
Q1	495	204	291	2.2e-16	A
Q2	464	228	236	2.2e-16	A
Q3	406	298	108	4.693e-05	A
Q4	336	365	29	0.273	N
	1 701	1 095	606	2.2e-16	A

**Tabulka W: Po–Pt (Q1–4)**

Je z ní patrné, že varianta dotazníku Q4 se vymyká vzorci ostatních variant a bude potřeba se jí věnovat blíže (viz Jednotlivé kontexty, 4.3.3). Pokud bychom opět hledali podobnosti mezi jednotlivými variantami dotazníku, mohli bychom vyčlenit variantu Q4 do jedné skupiny, v níž je rozdíl mezi hodnotami nejnižší, jádrem druhé skupiny, kde jsou rozdíly dobře patrné, by byla varianta Q1. Varianty Q2 a Q3 bychom museli zařadit mezi tyto dvě skupiny, neboť data, která z nich vycházejí, nejsou přesvědčivá ani ve prospěch hodnoty Po, ani ve prospěch hodnoty Pt.

Tabulka X zachycující naměřené hodnoty v závislosti na variantě lexikálního obsazení v dané verzi dotazníku ukazuje, že varianta dotazníku Q4 je velice odlišná od ostatních tří variant. Zároveň je zřejmé, že nacházíme tendenci k navazování na podmět

i v případě, že podle naší hypotézy by preferovanou variantou navazování koreferenčního řetězce měla být varianta předmětná. O tom svědčí zejména sloupec  $P_t > P_o$ , který zachycuje kontext, v němž je nastaveno předmětné uspořádání, ale respondenti přesto za antecedent zájmena označují podmět předchozí výpovědi, jedná se o kontexty typu *Fotbalisty Bohemians byli na zápas pozváni kluci. Holky je obdivovaly*. Tomu nasvědčuje i průměr počtu měření, v nichž respondenti považují za antecedent zájmena podmět (425.2) ve srovnání s tím, v kolika měřeních průměrně považují za antecedent předmět předchozí věty (273.8). Rozdíl mezi těmito průměry je větší než v případě topiku (385.5) a fokusu (317).

	$P_o > P_o$	$P_o > P_t$	$P_t > P_t$	$P_t > P_o$
Q1	224	129	75	271
Q2	236	109	119	228
Q3	212	141	157	194
Q4	100	251	114	236

**Tabulka X: Po–Pt predikce**

Při srovnání dat týkajících se aktuálního členění a dat týkajících se větněčlenské platnosti dojdeme k dílčím závěrům:

- varianty výchozího uspořádání T a  $P_o$ , přestože se shodují (formálně jde o stejnou variantu kontextu, konkrétně T– $P_o$ ), vykazují rozdílné hodnoty, tj. počty respondentů, kteří volí za antecedent koreferenčního řetězce T, nejsou stejné s počty respondentů, kteří volí za antecedent koreferenčního řetězce  $P_o$ . Tato disproporce závisí na tom, s jakým uspořádáním ji porovnáváme (F,  $P_t$ ), z čehož lze usuzovat, že tendence k navazování v rámci proměnné aktuálního členění a proměnné větněčlenské platnosti neovlivňují navazování koreferenčních řetězců ve stejné míře;
- hypotéza o preferovaném navazování na topik predikuje navazování koreferenčního řetězce s menší mírou spolehlivosti než hypotéza o preferovaném navazování na podmět, přesto stále na úrovni konvenční devadesátipětiprocentní hranice spolehlivosti.

#### 4.3.3 Jednotlivé kontexty

Klíčovým faktorem, který potenciálně ovlivňuje navazování koreferenčních řetězců, je lexikální obsazení jednotlivých položek. Každé lexikální obsazení totiž implikuje určitou komunikační situaci, na jejímž základě dochází k interpretaci koreferenčních řetězců v našem případě adresátem textu. Proto se nyní budeme zabývat hodnocením jednotlivých kontextů z hlediska odpovědí respondentů v dotazníkovém šetření. Musíme zohlednit nejen potenciální vliv lexikálního obsazení, ale i rozdíly mezi jednotlivými variantami jednoho lexikálního obsazení.

Následující tabulka Y zachycuje u každého kontextu naměřené hodnoty v nastavení topikovém, resp. podmětovém (první řádky) a hodnoty v nastavení fokusovém, resp. předmětovém (druhé řádky). Číslování kontextů odpovídá jejich uspořádání v příloze Sledované položky (10.4). Na základě hypotéz o převaze navazování na T a Po předpokládáme, že by hodnoty signalizující navazování na T v topikovém nastavení a hodnoty signalizující navazování na Po v podmětovém nastavení měly být větší než hodnoty signalizující navazování na F ve fokusovém nastavení a na Pt v předmětném nastavení. V příslušných prvních řádcích sloupců T a Po by tedy měly být hodnoty vyšší než hodnoty v příslušných prvních řádcích sloupců F a Pt. Z toho vyplývá i předpoklad týkající se nastavení fokusového a předmětného, tedy příslušných druhých řádků. V nich by naopak hodnoty měly být vyšší ve sloupcích F a Pt, protože předpokládáme, že se změnou nastavení topikového na fokusové se promění i navazování z původního topikového na fokusové<sup>86</sup>. Nejvyšší hodnoty by tedy podle testovaných hypotéz měly být v tabulce v prvních řádcích sloupce T a druhých řádcích sloupce F. Stejně pravidlo platí i pro větněčlenskou distinkci: nejvyšší hodnoty by měly být v prvních řádcích sloupce Po a druhých řádcích sloupce Pt.

kontexty	T	F	T–F	Po	Pt	Po–Pt
1	65 5	14 55	$p < 0.05$	62 8	20 49	$p < 0.05$
2	60 11	10 58	$p < 0.05$	56 15	45 25	$p = 0.083$

<sup>86</sup> Původním topikovým, resp. fokusovým, navazováním neodkazujeme na navazování na T, nebo na F, ale k navazování na jmenné fráze N1 a N2 v topikovém nebo fokusovém nastavení kontextu.

3	25 46	32 38	$p = 0.272$	32 39	61 9	$p < 0.05$
4	21 49	39 28	$p < 0.05$	25 45	58 12	$p < 0.05$
5	37 34	48 23	$p = 0.87$	37 34	52 19	$p < 0.05$
6	68 3	2 69	$p < 0.05$	64 3	10 59	$p < 0.05$
7	40 31	34 35	$p = 0.504$	46 23	44 27	$p = 0.687$
8	65 6	10 61	$p < 0.05$	62 8	39 32	$p < 0.05$
9	45 26	15 56	$p < 0.05$	47 22	54 24	$p = 1.000$
10	7 63	52 19	$p < 0.05$	17 53	56 15	$p < 0.05$
11	50 21	13 57	$p < 0.05$	54 17	57 13	$p = 0.566$
12	45 26	40 30	$p = 0.559$	45 25	40 28	$p = 0.628$
13	16 53	54 17	$p < 0.05$	12 58	60 10	$p < 0.05$
14	66 5	2 69	$p < 0.05$	65 6	17 54	$p < 0.05$
15	45 26	29 42	$p < 0.05$	48 23	54 14	$p = 0.167$
16	43 27	25 44	$p < 0.05$	39 31	53 13	$p < 0.05$
17	23 48	37 34	$p < 0.05$	21 49	41 28	$p < 0.05$
18	14 54	56 14	$p < 0.05$	10 60	52 17	$p < 0.05$
19	10 59	62 8	$p < 0.05$	10 60	64 7	$p < 0.05$
20	22 49	17 53	$p = 0.483$	20 51	61 10	$p < 0.05$

**Tabulka Y: kontexty chi-square (p-value)**

Při pohledu na tabulku Y je zřejmé, že naši hypotézu o tendenci k navazování na T a Po nepotvrzují všechny kontexty, dvacet sledovaných položek můžeme z tohoto hlediska uspořádat do tří základních skupin, které zachycuje následující tabulka Z.

skupina	počet	kontexty
I - potvrzující hypotézu	7	1, 2, 6, 7, 8, 12, 14



II - nepotvrzující hypotézu	8	3, 4, 5, 10, 13, 17, 18, 19
III - AČV potvrzující, VČ nepotvrzující hypotézu	5	9, 11, 15, 16, 20

**Tabulka Z: kontexty (ne)potvrzující hypotézy**

První skupinu tvoří sedm kontextů, v nichž v rámci obou proměnných (aktuálního členění i větněčlenské platnosti) nacházíme hodnoty potvrzující správnost hypotéz o navazování na T a Po. Zároveň v případě čtyř z nich (kontextů 1, 6, 8, 14) jsou hodnoty signifikantní u obou proměnných, jak ukazuje tabulka Y, u jednoho (kontext 2) z nich je signifikantní jen rozložení hodnot v rámci AČV, u zbylých (kontexty 7 a 12) není rozložení hodnot signifikantní ani v rámci proměnné AČV, ani v rámci VČ.

Druhou skupinu tvoří osm kontextů, které v rámci dotazníkového výzkumu hypotézy o převažujícím navazování na T a Po nepotvrdily. Jedná se o kontexty, kdy jsou hodnoty v rámci obou proměnných signifikantní, ale v rozporu s hypotézami (kontexty 4, 10, 13, 17, 18, 19), a o kontexty, kdy je signifikantní pouze rozložení hodnot v rámci VČ, nikoli v rámci AČV (kontexty 3, 5).

Třetí skupinu tvoří kontexty, u nichž nebyla tendence vykazovaná v rámci proměnných konzistentní. Všechny kontexty, které jsme zařadili do této skupiny, v rámci měření proměnné AČV hypotézy potvrzují, zatímco v rámci měření proměnné VČ hypotézu nepotvrzují. Nutno ale dodat, že v této skupině je nejmenší míra signifikantnosti výsledků, signifikantní v rámci měření obou proměnných je pouze jeden kontext (kontext 16), signifikantní jen v rámci potvrzení hypotézy týkající se AČV jsou tři kontexty (kontexty 9, 11 a 15) a signifikantní jen v rámci nepotvrzení hypotézy týkající se VČ je jeden kontext (kontext 20).

Zajímavé je, že ve skupině I, tedy mezi kontexty potvrzujícími hypotézy, nenacházíme takový typ hodnot, který by signifikantně potvrdil tendenci k navazování na podmět, ale zároveň vykazoval nesignifikantní hodnoty v rámci proměnné AČV, které by hypotézu také potvrzovaly. V případě skupiny II, tedy mezi kontexty nepotvrzujícími hypotézy, je situace přesně opačná, nenacházíme kontexty, které by vykazovaly signifikantní hodnoty v případě AČV a zároveň vykazovaly nesignifikantní hodnoty v případě VČ. Kromě toho v rámci kontextů II nepotvrzujících hypotézy nenacházíme ani hodnoty, které by v rámci obou proměnných vykazovaly nesignifikantní hodnoty. I v rámci skupiny, v níž se hodnoty týkající se jednotlivých proměnných nechovají navzájem

konzistentně, jedna logicky možná varianta nenachází v datech opodstatnění. Všechny kontexty ve skupině III potvrzují hypotézu týkající se AČV, ale nepotvrzují hypotézu týkající se VČ, nenacházíme hodnoty, které by potvrzovaly tendenci v rámci VČ, ale nepotvrzovaly tendenci v rámci AČV.

Kontexty, které nepotvrzují hypotézy (II), z velké většiny v hodnocení přijatelnosti patřily do skupiny, která byla hodnocena pozitivně (kontexty 4, 10, 13, 18, 19), jen malá část z nich jsou kontexty nově vytvořené namísto vyřazených negativně hodnocených kontextů (kontexty 3 a 17) nebo upravené (kontext 5). To, že byly v rámci testování přijatelnosti kontexty hodnoceny pozitivně, vylučuje vliv uměle vytvořeného kontextu na vlastní měření. Disproporce mezi naprosto přijatelnými kontexty a tím, že nepotvrzují hypotézy o navazování koreferenčních řetězců, považujeme za vliv lexikálního obsazení.

Skupina kontextů potvrzujících hypotézy (I) nevykazuje takovou převahu kontextů hodnocených v rámci testování přijatelnosti podobně. Nacházíme kontexty, které byly hodnoceny pozitivně (kontexty 2, 8, 14), kontexty upravené (kontexty 1, 7) i kontexty nově vytvořené (kontexty 6, 12).

Kontexty ve skupině III bychom z hlediska výsledků testování přijatelnosti mohli rozdělit téměř na poloviny: kontexty hodnocené pozitivně (kontexty 11, 16) a negativně (kontexty 9, 15, 20). Podrobnější přehled poskytne srovnání s tabulkou N zachycující výsledky testování přijatelnosti.

Zajímavý je fakt, že všechny kontexty, které byly v rámci testování přijatelnosti hodnoceny negativně, patří právě do skupiny III, zatímco všechny kontexty vytvořené nově namísto kontextů hodnocených negativně patří buď do skupiny I nebo II. Tyto korelace by ovšem bylo potřeba podrobněji prozkoumat, neboť není možné na základě vzorku o rozsahu dvaceti kontextů spolehlivě stanovit vazby mezi přijatelností (resp. jejím hodnocením) a (ne)potvrzením našich pracovních hypotéz.

#### 4.3.4 *Zájmena*

Mohli bychom předpokládat, že to, jaký antecedent je zájmenu přiřazen, je závislé nejen na aktuálněčlenském (slovosledném) a větněčlenském uspořádání předcházející výpovědi, ale také na zájmenu samotném. Aby byl tento případný vliv typu zájmena na navazování koreferenčního řetězce eliminován, bylo do testovacích kontextů začleněno více typů zájmen (viz *Zájmenná koreference*, 4.2.5). Vzhledem k získaným výsledkům

představeným výše nicméně nebyl prokázán vliv typu zájmena na to, jaký antecedent je mu přiřazen, s jakou částí předchozí věty navazuje koreferenční řetězec.

Ve třinácti kontextech se zájmeny osobními (1–4, 8, 9, 11, 14, 15, 17–20) existuje preference k navazování na T stejně jako v případě sedmi kontextů se zájmeny přivlastňovacími (5–7, 10, 12, 13, 16) (viz tabulka L s označením typu zájmena a tabulka Y s naměřenými hodnotami podle jednotlivých kontextů).

V případě aktuálněčlenské proměnné je tendence k navazování na T silnější v případě osobních zájmen téměř dvakrát (1 825:985, 1.85) a tento rozdíl je signifikantní ( $X^2$  test: chi-squared = 22.32, df = 1, p-value = 2.313e-06). Z toho vyplývá, že tendenci k navazování na T vykazují zájmena osobní silněji než zájmena přivlastňovací, což pravděpodobně souvisí s tím, že osobní zájmena častěji stojí v podmětu než zájmena přivlastňovací, a proto se u nich projevuje tendence k přechodu typu continue zachovávajícímu podmět silněji než u zájmen přivlastňovacích.

V případě větněčlenské proměnné není mezi osobními a přivlastňovacími zájmeny signifikantní rozdíl, zajímavý je z hlediska větněčlenské proměnné kontext 9 (*Pan domácí srdečně přivítal Karla. Půjčil mu peníze.*), který jako jediný vykazuje preferenci navazování na Pt v podmětovém i předmětovém nastavení. Všechny ostatní kontexty, bez ohledu na typ zájmena v nich obsaženého, vykazují variaci navazování koreferenčního řetězce v závislosti na variantě kontextu (podmětové vs. předmětové nastavení), ať už ve prospěch potvrzení hypotézy, nebo ve prospěch vyvrácení hypotézy.

#### 4.3.5 Kritická místa výzkumu

Zásadní vliv na získané výsledky může mít velké množství faktorů. Zmíním proto jen některé. Tím nejdůležitějším je zcela jistě sestavení dotazníkových položek do jednotlivých variant dotazníku. Pokud není možné využívat ve výzkumu autentický jazykový materiál, je naprosto nutné provést u uměle vytvořeného materiálu předvýzkum, v našem případě testování přijatelnosti kontextů. Je samozřejmé, že ani tímto testováním není zaručeno kompletní odstínění vlivu umělosti testovacích položek na výsledky, a že nutnost provedení výzkumu přijatelnosti přináší zároveň riziko zkreslení výsledků a metodologických chyb. Zároveň je ale vliv uměle vytvořeného jazykového materiálu testováním přijatelnosti do značné míry eliminován.

Druhým kritickým místem představeného dotazníkového šetření je čas, který je respondentům dán k dispozici v dotazníkovém šetření. Nastavení dvaceti sekund jako limitu bylo založeno na testu několika potenciálních respondentů tak, aby nebyl respondentům poskytnut zbytečně velký prostor, který by jim umožňoval přemýšlet nad odpovědí, ale zároveň tak, aby stihli pohodlně přečíst veškerý potřebný text a vybrat variantu odpovědi. Jsme si vědomi toho, že časový limit může respondenta ovlivňovat, signalizace zbývajících času může působit rušivě při čtení varianty kontextu i možností odpovědi a už jen vědomí toho, že je respondent limitován, může ovlivnit jeho odpovědi. Z celkového počtu respondentů jich ale jen minimum po ukončení testu vzneslo námitku na příliš omezující časový limit, proto jsem toho názoru, že dvacetivteřinový časový limit byl zvolen vhodně. Přes všechny potenciálně negativní důsledky si myslím, že určení časového limitu je vhodným nástrojem, který poskytuje možnost se alespoň trochu přiblížit primárnímu čtení bez nutnosti upozorňovat na nejednoznačnost navazování koreferenčního řetězce.

Třetím a posledním zásadním kritériem, které zmíníme, je interpretace výsledků dotazníkového šetření a dat získaných z PDT. Je zřejmé, že zvolená metodologická perspektiva (teorie centeringu, Danešův koncept tematických posloupností) výrazně ovlivní způsob interpretace výsledků. Přijetí metodologického přístupu je pro práci s jazykovými daty nutné, důležité proto je být si vědom omezení, která jsou s přijatou metodologií spojena (viz kapitola Centering, 3.4).

## 5 Závěr

Nyní shrneme cíl, metody, postup a závěry, k nimž jsme došli. Věnovali jsme se zkoumání navazování koreferenčních řetězců na základě aktuálního členění a větněčlenské platnosti. K oběma přistupujeme jako k proměnným, které ovlivňují utváření koreferenčních vazeb v textu. Zkoumány byly následující hypotézy: a) hypotéza o převažujícím navazování na topikovou část předchozí výpovědi, b) hypotéza o převažujícím navazování na podmět předcházející výpovědi.

Představili jsme teorii aktuálního členění tak, jak ji známe z českého prostředí, včetně konceptu tematických posloupností (Mathesius, Daneš), metodologii přístupu ke koreferenci obecně i v rámci PDT a vztah větněčlenské platnosti ke koreferenci (Zimová). Kromě toho jsme se věnovali představení tzv. teorie centeringu. Mathesiův přístup byl zvolen proto, že je považován za zakladatele české teorie aktuálního členění, Danešův proto, že jeho přístup je přínosem ke zkoumání tematických posloupností. Zimová poskytuje metodologický základ pro zkoumání proměnné větněčlenské platnosti, její práce byla zvolena z toho důvodu, že propojuje perspektivu výstavby textu z hlediska navazování s perspektivou větněčlenskou. Teorie centeringu je pro naše zkoumání vhodným zdrojem, protože poskytuje predikce o navazování koreferenčních vztahů v textu a zároveň doplňuje Danešovu perspektivu tematických posloupností.

Z těchto přístupů vychází hypotézy o tom, že navazování koreferenčního řetězce probíhá častěji na T než na F, a o tom, že navazování koreferenčního řetězce probíhá častěji na Po než na Pt. Obě tyto hypotézy byly ověřovány na dvou typech dat popsanych v následující části.

### 5.1 *Utváření koreferenčních řetězců recipientem textu*

Jako hlavní zdroj dat pro zkoumání koreference posloužilo měření koreferenčních vazeb v PDT a dotazníkový výzkum. V PDT byla na materiálu 2064 koreferenčních vztahů zjištěna a) preference navazování z topikové části výpovědi (2031:33, 61.55), b) preference navazování do topikové části výpovědi předcházející (1199:865, 1.39), c) v případě větněčlenské platnosti preference navazování z předmětu aktuální výpovědi (480:578, 0.83) a d) preference navazování na podmět předcházející výpovědi (942:708, 1.33).

Všechny uvedené hodnoty jsou signifikantní na devadesátipětiprocentní hladině spolehlivosti.

Dotazníkové šetření, kterého se účastnilo 284 respondentů, bylo zaměřeno na to, jakým způsobem budou recipienti textu navazovat koreferenční řetězce v textu, pokud bude kontext nejednoznačný (typu *Populární zpěvačka byla oceněna cenou Grammy. Fanoušci jsou z ní nadšení.*). Každému recipientovi bylo předloženo dvacet sledovaných a třicet výplňkových kontextů, u nichž bylo jeho úkolem určit koreferenční vztahy, přičemž zobrazení položek bylo časově omezeno, aby respondenti odpovídali podle svého preferovaného čtení. Struktura kontextů v dotazníku variovala na základě kombinace postavení jmenných frází v topiku/fokusu a v podmětu/předmětu, aby bylo možné sledovat, zda se se změnou postavení aktuálněčlenského a větněčlenského promění určení antecedentu. Každý kontext tedy existuje ve třech variantách, z nichž jedna je topiková a zároveň podmětová, např. *Liberální voliči<sub>T,Po</sub> podporují demokratické kandidáty<sub>F,Pt</sub>. Bojují za občanské svobody, volný trh a zmírnění chudoby.*, a k nim přísluší další varianty, v nichž vždy variuje jen hodnota jedné z proměnných: jedna varianta fokusová, v níž se oproti variantě topikové promění postavení v T/F, ale zachováno zůstává rozvržení větněčlenských rolí na Po a Pt, např. *Demokratické kandidáty<sub>T,Pt</sub> podporují liberální voliči<sub>F,Po</sub>. Bojují za občanské svobody, volný trh a zmírnění chudoby.* a jedna varianta předmětová, v níž variuje větněčlenská příslušnost jmenných frází k Po a Pt, ale rozvržení v T a F zůstává zachováno, např. *Liberálními voliči<sub>T,Pt</sub> jsou podporováni demokratičtí kandidáti<sub>F,Po</sub>. Bojují za občanské svobody, volný trh a zmírnění chudoby.* Na základě takto navrženého výzkumu jsme získali tyto výsledky:

- konkrétní lexikální obsazení je pro navazování koreferenčních řetězců klíčové, neboť s ním je asociována konkrétní komunikační situace, která může vykazovat preference ovlivňující navazování koreferenčních řetězců,
- čeština, jakožto jazyk s volným slovosledem, nevykazuje tak přesvědčivé rozdíly mezi jednotlivými variantami kontextů ve smyslu vlivu varianty na způsob navazování koreferenčních řetězců, jako je tomu v případě angličtiny,
- preferovaným navazováním je odkazování na topik předchozí výpovědi (1542:1268, 1.22),

- z hlediska větněčlenského je dominantním navazování na podmět předchozí výpovědi (1701:1095, 1.55),
- podmět je jako antecedent volen i v případech, kdy kontext má nastavení předmětové a na základě hypotézy o preferenci navazování na topik by měl být za antecedent označován předmět (tabulka X a oddíl Síla proměnných, 5.4),
- jedna z variant dotazníku, Q4, vykazuje poněkud odlišné hodnoty než varianty ostatní, v případě aktuálního členění potvrzuje hypotézu o navazování na T spíše než na F, ale velice slabě (363:344, 1.06), v případě větněčlenské platnosti jsou hodnoty přímo opačné než u ostatních variant, svědčí ve prospěch navazování na Pt spíše než na Po (336:265, 1.92), tyto hodnoty ovšem nejsou signifikantní, tudíž mohou být pouze náhodné.

Největší počet vztahů z topiku do topiku svědčí o tom, že preferovaným schématem výstavby textu je tematická posloupnost typu průběžné téma; hypotéza o navazování na topikovou část výpovědi byla potvrzena, i když nijak výrazným rozdílem. Hypotéza o navazování na podmět založená na centeringových predikcích preference typu přechodu continue, případně retain, byla potvrzena také.

## **5.2 Centeringová teorie a čeština**

Teorie centeringu pracuje s členěním výpovědní struktury na tzv. centra pozornosti tří typů: 1. Cf – centra pozornosti, která zakládají možnost, aby na ně bylo odkazováno jako na antecedent, 2. Cp – nejvýše postavené centrum Cf, tj. v češtině prakticky vždy podmět aktuální výpovědi, 3. Cb – centrum pozornosti, které se vyskytuje v aktuální výpovědi a zároveň se vyskytuje i ve výpovědi předcházející, konkrétně jde o Cp předcházející výpovědi. Na základě vztahů mezi Cb a Cp aktuální výpovědi a vztahů mezi Cb aktuální a předcházející výpovědi člení teorie centeringu posuny, tj. principy výstavby textu, do čtyř typů: 1. continue:  $Cb(U_i) = Cp(U_i)$ ,  $Cb(U_i) = Cb(U_{i-1})$ , 2. retain:  $Cb(U_i) \neq Cp(U_i)$ ,  $Cb(U_i) \neq Cb(U_{i-1})$ , 3. smooth-shift:  $Cb(U_i) = Cp(U_i)$ ,  $Cb(U_i) \neq Cb(U_{i-1})$ , 4. rough-shift:  $Cb(U_i) \neq Cp(U_i)$ ,  $Cb(U_i) \neq Cb(U_{i-1})$ .

Aplikace teorie centeringu na český materiál se ukázala jako možná, nicméně ne zcela soběstačná metoda. Omezení centeringové teorie pouze na lokální, tj. bezprostřední,

koreferenci se ukázalo jako bezproblémové, neboť většina koreferenčních řetězců je navazována lokálně.

Nicméně tendence, které centeringová teorie predikuje, nejsou na českém materiálu doložitelné v takové míře. V češtině se zdá být mnohem častější navazování na fokusovou část výpovědi, než je tomu v angličtině, tedy, zjednodušeně řečeno, v češtině se více než v angličtině vyskytuje tematická posloupnost typu tematizace rématu a centeringový přechod typu shift, které nelze ztotožňovat, ale které mají společné např. to, že topik (zjednodušeně můžeme topik připodobnit k centru pozornosti Cb) není v rámci předcházející výpovědi a výpovědi aktuální totožný. Data získaná dotazníkovým výzkumem potvrzují preferenci typu přechodu continue, navazování na podmět předchozí výpovědi. Zároveň ale data z PDT ukazují jako dominantnější typ přechodu retain (antecedentem koreferenčního vztahu je podmět předchozí výpovědi, anaforem předmět výpovědi aktuální), ačkoli centeringová teorie preferuje typ continue (antecedentem i anaforem jsou podměty příslušných výpovědí). Continue i retain patří podle centeringové teorie k těm častějším způsobům přechodů, zkoumání toho, který z nich je častější v češtině, by vyžadovalo hlubší analýzu i z hlediska typu textu (mohli bychom předpokládat, že v různých typech textu, například v textech spontánních vs. redigovaných, probíhá navazování koreferenčních řetězců jiným způsobem).

### **5.3 Referenční nespecifikovanost**

Nejednoznačnost kontextů v rámci dotazníkového výzkumu byla uměle vytvořena zejména proto, že od každého lexikálního obsazení musely být vytvořeny tři varianty na základě změny postavení jedné jmenné fráze v topiku a fokusu, resp. podmětu a předmětu. V autentickém českém textu je běžné, že nejsou vztahy na jazykové rovině jednoznačné. Referenční nespecifikovanost, tj. nejednoznačná volba antecedentu především zájmena, je pro běžnou komunikaci typická, ke zjednoznačnění reference dochází na úrovni mimojazykové (ostenzí, znalostí mimojazykového kontextu apod.). Svědčí o tom například případy zachycené v autentických jazykových vyjádřeních: *Takhle si boty nemůžu obout, musím si umýt nohy. Jsou moc špinavé.* [soukromá konverzace, 22. 8. 2012], *Herec vyslovuje myšlenky vymyšlené někým jiným. Někdy ani ne příliš zdařile.* [Lidové noviny, 25. 8. 2012], *To rozložení sil, že on je prezident a on je ministr zahraničí, že je velice dobré rozložení sil.* [televizní pořad 100 dní na Hradě – ohlédnutí, 11. 6. 2013]. Ostatně jeden



z autentických kontextů byl zařazen do výzkumu jako kontrolní položka, zda se respondenti věnují dotazníku pozorně, neboť v tomto případě existovala mimojazyková znalost, která určovala „správnou“ variantu volby antecedentu: *Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země.* A ukázalo se, že tato mimojazyková znalost je naprosto klíčová – chybnou volbu antecedentu jsme zaznamenali jen u pěti respondentů, 137 z nich zvolilo správně.

Sidner (1981: 225) také připouští, že nejednoznačnost kontextu je přirozená, nicméně on k vysvětlení přistupuje jinak. Na základě kontextu *Sally si koupila sekačku, která měla rozbitou čepel. Viděla na ni reklamu v televizi.* ukazuje, že reklama v televizi pravděpodobně neukazovala sekačku s rozbitou čepelí a koreferenční řetězec není navázán na *sekačku, která měla rozbitou čepel*, odkazuje k typu, ne ke konkrétnímu reprezentantovi skupiny *sekaček*. V tom podle Sidnera spočívá ambiguita tohoto kontextu. Odkazování tohoto typu je velice časté.

#### **5.4 Síla proměnných**

Pokud jde o vztah aktuálního členění a větněčlenské struktury, může nám k jeho posouzení posloužit například aplikace akvizičního kritéria. Deszó-Szépe tvrdí, že „základní pravidla týkající se topiku a commentu se zdají být osvojována dříve, než nastává formování takových kategorií povrchové struktury jako podmět a předmět“.<sup>87</sup> V důsledku tohoto tvrzení bychom za primární proměnnou mohli považovat aktuální členění.

Při srovnání, jak silný vliv na navazování koreferenčních řetězců mají zkoumané proměnné v rámci dotazníkového šetření, tedy podle hodnocení rodilých mluvčích češtiny, pozorujeme, že tendence k navazování na podmět je silnější než tendence k navazování na topik (1.22:1.55), což se projevuje v případech, kdy je podmět umístěn ve fokusu a předmět v topiku (např. *Nového generálního ředitele Alexandra Schramma jmenoval letos Ronald McDonald. Jeho prozřetelnost zachránila firmě reputaci.*) a podle první hypotézy zahrnující aktuálněčlenskou proměnnou by měl respondent zvolit za antecedent předmět v topiku, ale přesto volí také podmět ve fokusu. Tato disproporce je dobře patrná v tabulce Y.

<sup>87</sup> „The primary topic – comment rules seem to occur earlier in the acquisition of language than the formativ of such categories of surface structure as Subject and Object.“ (Deszó-Szépe 1974: 85)

### **5.5 Slabiny dosavadních přístupů**

Aplikované přístupy, jak teorie Danešových tematických posloupností, tak teorie centeringu, mají jistá omezení, s nimiž je při výzkumu nutno počítat. Danešovy tematické posloupnosti představují ideální konstrukt a není často jednoznačné, jaký typ tematické posloupnosti určit, neboť se částečně překrývají. Kromě toho nepracuje Danešova metoda určování typu tematické posloupnosti s větněčlenskou platností jednotlivých koreferovaných členů, která ale má, jak jsme ukázali, na utváření koreferenčních vztahů zásadní vliv. Teorie centeringu je omezena pouze na modelování lokální koreference, což se neukázalo jako problém v případě našeho výzkumu, ale pokud bychom měli ambice popsat všechny koreferenční vztahy, ne jen velice specificky vymezenou skupinu, tato metodologie by byla nedostačující.

## 6 Seznam literatury

ČERMÁK, František (2010): *Lexikon a sémantika*. Praha: NLN.

DANEŠ, František (1968): Typy tematických posloupností v textu (na materiále českého textu odborného). *SaS* 29, s. 125–141.

DANEŠ, František (1974a): Functional Sentence Perspective and the Organization of the Text. In *Papers on functional sentence perspective*. Daneš (ed.). Praha: Academia. Vydání první.

DANEŠ, František et al. (1974b): Zur Terminologie der FSP. In *Papers on functional sentence perspective*<sup>88</sup>. Daneš (ed.). Praha: Academia. Vydání první.

DANEŠ, František – HLAVSA, Zdeněk – KOŘENSKÝ, Jan (1973): *Práce o sémantické struktuře věty*. Praha: ÚJČ AV ČSAV. Vydání první.

DEZSÖ, László – SZÉPE, Görgy (1974): Two problems of topic – comment. In *Papers on Functional Sentence Perspective*. Daneš (ed.). Praha: Academia.

DOKULIL, Miloš (1958): K povaze vztahu slova a pojmu, věty a myšlenky. In *O vědeckém poznání soudobých jazyků*. Dostál (ed.) Praha: Nakladatelství Československé akademie věd.

DVOŘÁČKOVÁ, Eva (2005): *Topic-focus articulation and the patterning of discourse*. Praha: nepublikovaná diplomová práce (vedoucí Eva Hajičová).

Encyklopedický slovník češtiny (2002). Praha: NLN.

ERTESCHIK-SHIR, Nomi (2007): *Information structure: the syntax–discourse interface*. Oxford: Oxford University Press.

FIRBAS, Jan (1986): On the Dynamics of Written Communication in the Light of the Theory of Functional Sentence Perspective. In *Studying Writing: Linguistic Approaches*. Cooper, Greenbaum (eds.). Beverly Hills: SAGE, s. 40–71.

FIRBAS, Jan (1992): *Functional sentence perspective in written and spoken communication*. Cambridge: Cambridge University Press.

---

<sup>88</sup> Daneš 1974.

GREPL, Miroslav – KARLÍK, Petr (1986): *Skladba spisovné češtiny*. Praha: SPN.

GROSZ, Barbara – WEINSTEIN, Scott – JOSHI, Aravind (1995): Centering: A Framework for Modeling the Local Coherence of Discourse. In *Computational Linguistics* 2, s. 203–225.

HAJIČOVÁ, Eva (1999): Aktuální členění věty a výstavba promluvy. In *Čeština – univerzália a specifika* 1. Brno: ÚČJ FF MU, s. 47–54.

HAJIČOVÁ, Eva (2003): Contextual boundness and diskurse patterns. In *Proceedings of XVIIth International Congress of Linguists*. Hajičová, Kotěšovcová, Mírovský (eds.) Praha: ÚFAL MFF UK.

HAJIČOVÁ, Eva – PARTEE, Barbara – SGALL, Petr (1998): *Topic, focus articulation, tripartite structures and semantic content*. Dordrecht: Kluwer Academic Press.

HARWEG, Roland (1994): Is There Implicit Inclusion of Preceding Predicates in Anaphorics of Nominals or Pronominals? In *The syntax of sentence and text: a festschrift for František Daneš*. Čmejrková, Štícha (eds.). Amsterdam, Philadelphia: John Benjamins, s. 131–146.

HOFFMANNOVÁ, Jana (1983): *Sémantické a pragmatické aspekty koherence textu*. Praha: ÚJČ AV ČR.

KUČOVÁ, Lucie – VESELÁ, Kateřina – HAJIČOVÁ, Eva – HAVELKA, Jiří (2005): Topic-focus articulation and anaphoric relations: A corpus based probe. In *Proceedings of Discourse Domains and Information Structure workshop*. Heusinger, Umbach (eds.). Edinburgh: Heriot-Watt University.

LAMBRECHT, Knud (1994): *Information structure and sentence form: topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

MATHESIUŠ, Vilém (1947): O tak zvaném aktuálním členění větném. In *Čeština a obecný jazykozpyt*. Praha: Melantrich. (poprvé publikováno SaS 5/1939, s. 171–174)

MEREU, Lunella (2009) (ed.): *Information structure and its interfaces*. Berlin, New York: Mouton de Gruyter.

MIKULOVÁ, Marie et al. (2005): *Anotace na tektogramatické rovině Pražského závislostního korpusu. Anotátorská příručka*. Praha: ÚFAL MFF UK.

Mluvnice češtiny (1986) 2. díl – Tvarosloví. Praha: Academia

Mluvnice češtiny (1987) 3. díl – Skladba. Praha: Academia

NEDOLUZHKO, Anna – MÍROVSKÝ, Jiří – OCELÁK, Radek – PERGLER, Jiří (2009): Extended Coreferential Relations and Bridging Anaphora in the Prague Dependency Treebank. In: Proceedings of Discourse Anaphora and Anaphor Resolution Colloquium. Sobha, Branco, Mitkov (eds.). Goa (India): University Trier.

NĚDOLUŽKO, Anna (2011): Rozšířená textová koreference a asociační anafora. Praha: ÚFAL MFF UK.

NOVÁK, Pavel (1959): O prostředcích aktuálního členění. In Acta Universitatis Carolinae – Philologica 1, s. 9–15.

PANEVOVÁ, Jarmila (1978): K významové stavbě větného centra. SaS 39, s. 140–148.

PANEVOVÁ, Jarmila (2013): Elipsa, aneb co v korpusu nenajdeme. Jazykovědné sdružení ČR, 16. 5. 2013.

PEREGRIN, Jaroslav (1995): Topic-Focus Articulation as Generalized Quantification. In Proceedings of Focus and natural language processing, Bosch, van der Snadt (eds.). Heidelberg: IBM Deutschland.

PRASAD, Rashmi – DINESH, Nikhil – LEE, Alan – MILTSAKAKI, Eleni – ROBALDO, Livio – JOSHI, Aravind – WEBBER, Bonnie (2008): The Penn Discourse TreeBank 2.0. In Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC). Marrakech (Morocco): European Language Resources Association.

SGALL, Petr (1972): Základ a jádro, nebo presupozice a ohnisko? SaS 33, s. 160–163.

SGALL, Petr (1974): Zur Stellung der Thema-Rhema-Gliederung in der Sprachbeschreibung. In Papers on Functional Sentence Perspective. Daneš (ed.). Praha: Academia.

SGALL, Petr (1976): K obecným otázkám sémantiky věty. SaS 37, s. 184–194.

SGALL, Petr – HAJČOVÁ, Eva – PARTEE, Barbara (1998): Topic-focus articulation, tripartite structure and semantic content. Dordrecht: Kluwer Academic Publishers.

SIDNER, Candace (1981): Focusing for Interpretation of Pronouns. In *American Journal of Computational Linguistics* 7, č. 4, s. 217–231.

SVOBODA, Aleš (1974): On Two Communicative Dynamisms. In *Papers on Functional Sentence Perspective*. Daneš (ed.) Praha: Academia.

STEELE, Ross – THREADGOLD, Terry (1987): *Language Topics – Essays In Honour Of Michael Halliday*. Volume 1, 2. Amsterdam, Philadelphia: John Benjamins.

TRÁVNÍČEK, František (1961): O tak zvaném aktuálním členění větném. *SaS* 22, s. 163–171.

ÚJČ (1970): *Functional sentence perspective: Papers prepared for the Symposium held at Mariánské Lázně (Marienbad) on 12th – 14th October 1970*. Praha: Nakladatelství Československé akademie věd.

WALKER, Marilyn – JOSHI, Aravind – FRIEDMAN PRINCE, Ellen (1998) (eds.): *Centering Theory in Discourse*. Oxford: Clarendon Press. First Edition.

ZIMOVÁ, Ludmila (1987): Vyjadřování podmětu osobními zájmeny 1. a 2. osoby. *NŘ* 71, s. 10–22.

ZIMOVÁ, Ludmila (1990): Podmět "pronominální" nebo "pronominalizovaný", "nevyjádřený" nebo "elidovaný"? In *Sborník prací Pedagogické fakulty v Ústí nad Labem. Řada bohemistická*. Praha: SPN, s. 5–12.

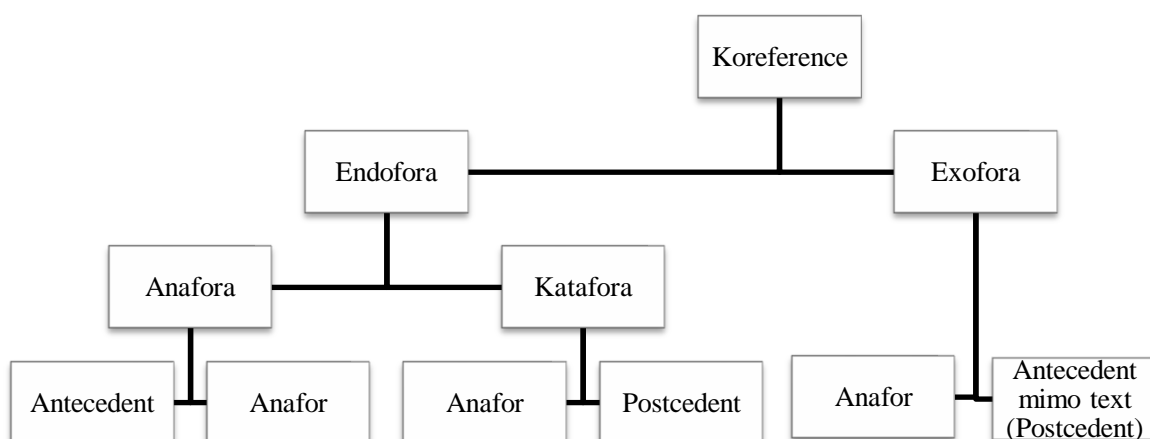
ZIMOVÁ, Ludmila (1991): Vyjadřování podmětu v textu. *NŘ* 74, s. 75–85.

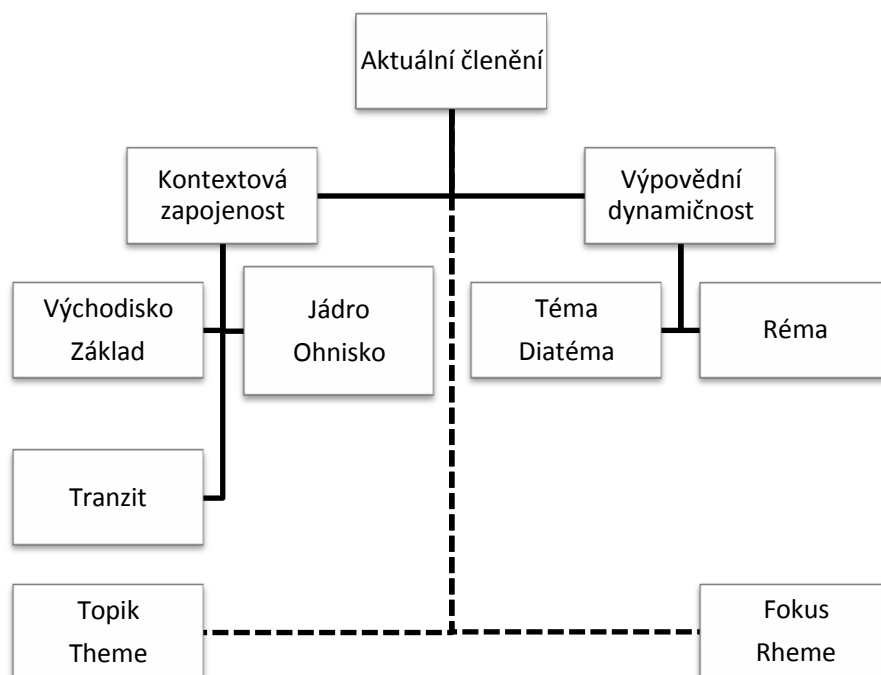
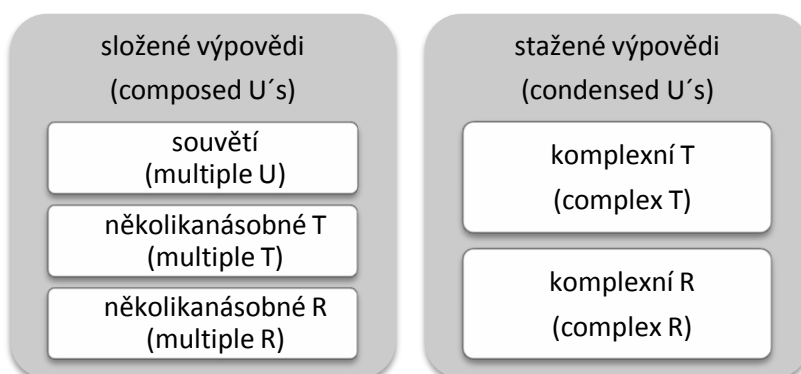
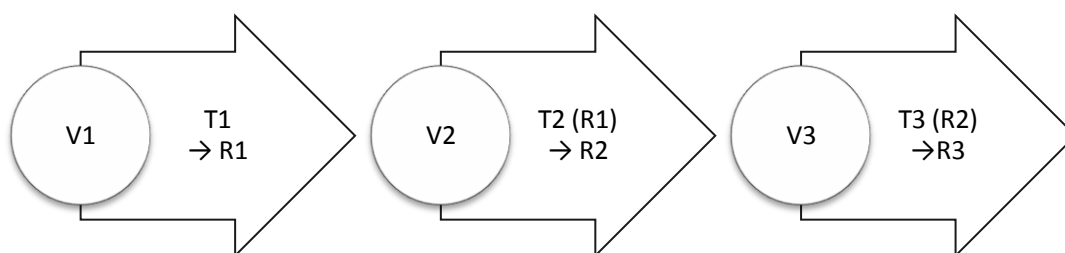
ZIMOVÁ, Ludmila (1994): Způsoby vyjadřování větných členů v textu: konkurence pojmenování, pronominalizace a elize. Ústí nad Labem: Univerzita J. E. Purkyně.

## 7 Seznam schémat

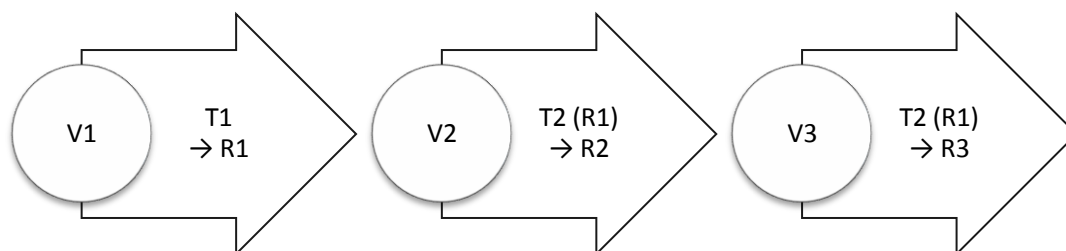
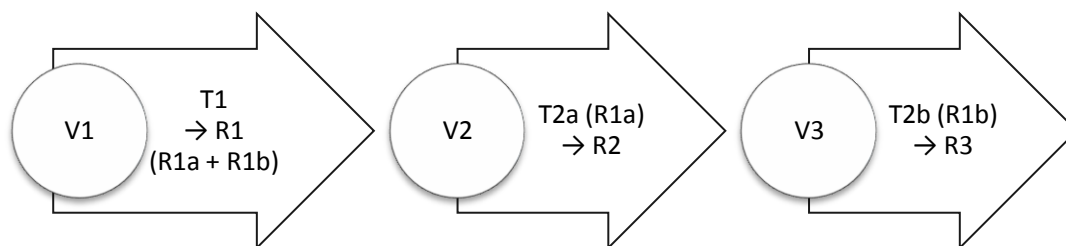
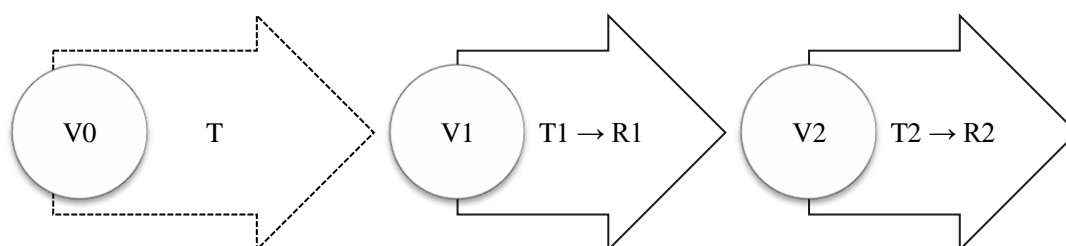
<i>Schéma A: koreference</i> .....	18
<i>Schéma B: aktuální členění</i> .....	22
<i>Schéma C: složené a stažené výpovědi</i> .....	36
<i>Schéma D: návazná tematizace rématu</i> .....	37
<i>Schéma E: průběžné téma</i> .....	38
<i>Schéma F: rozvíjení rozštěpeného rématu</i> .....	39
<i>Schéma G: tematické odvozování</i> .....	40
<i>Schéma H: tematický skok</i> .....	41
<i>Schéma I: anotace AČV v PDT</i> .....	62
<i>Schéma J: Specifičnost reference</i> .....	65

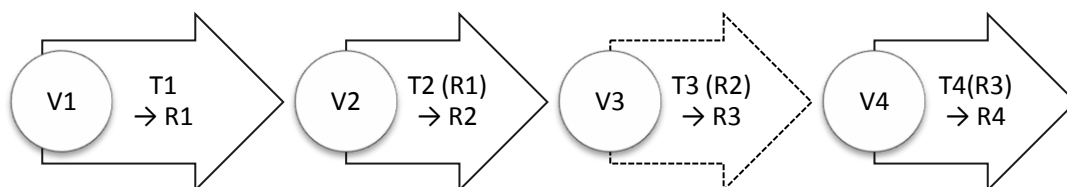
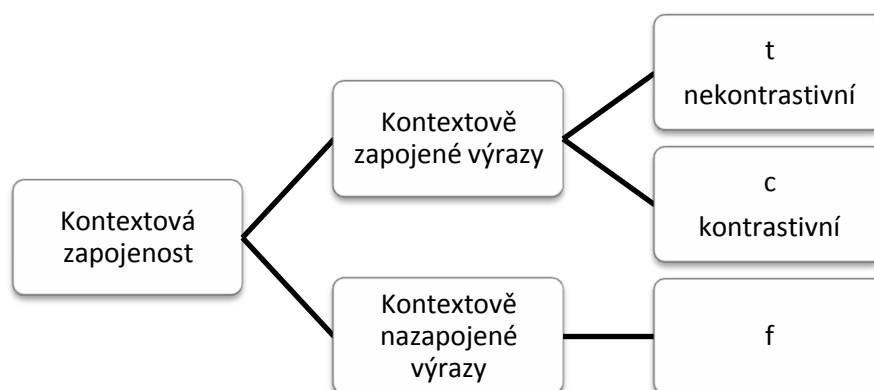
### 7.1 Schéma A: koreference

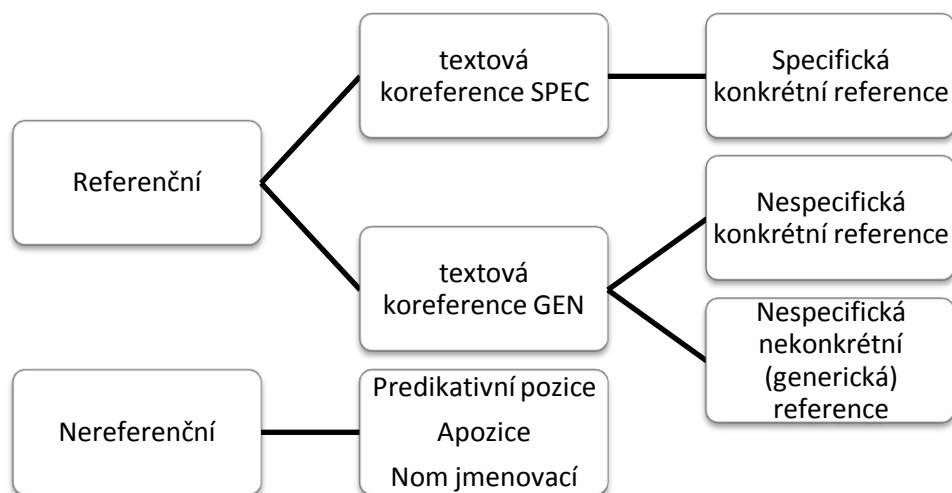


**7.2 Schéma B: aktuální členění****7.3 Schéma C: složené a stažené výpovědi****7.4 Schéma D: návazná tematizace rématu**



**7.5 Schéma E: průběžné téma****7.6 Schéma F: rozvíjení rozštěpeného rématu****7.7 Schéma G: tematické odvozování**

**7.8 Schéma H: tematický skok****7.9 Schéma I: anotace AČV v PDT**

**7.10 Schéma J: Specifičnost reference**

## 8 Seznam tabulek

<i>Tabulka A: formální a aktuální členění</i> .....	17
<i>Tabulka B: podmět, předmět</i> .....	48
<i>Tabulka C: „tematické“ posloupnosti v teorii centeringu</i> .....	54
<i>Tabulka D: Centering – analýza</i> .....	55
<i>Tabulka E: textová koreference (tfa)</i> .....	68
<i>Tabulka F: koreference v ef. br. (AVČ)</i> .....	70
<i>Tabulka G: koreference v ef. br. (VČ)</i> .....	71
<i>Tabulka H: t/c, f + Po, Pt v ef. br.</i> .....	72
<i>Tabulka I: přehled výsledků měření (PDT)</i> .....	73
<i>Tabulka J: schéma testovacích kontextů</i> .....	80
<i>Tabulka K: testovací kontext s notací</i> .....	81
<i>Tabulka L: sledované položky a zájmena</i> .....	85
<i>Tabulka M: kódování kontextů (přijatelnost)</i> .....	90
<i>Tabulka N: přijatelnost (akce)</i> .....	91
<i>Tabulka O: upravené kontexty</i> .....	92
<i>Tabulka P: nové kontexty</i> .....	93
<i>Tabulka Q: Q1–4</i> .....	98
<i>Tabulka R: nezaznamenané hodnoty</i> .....	105
<i>Tabulka S: T–F sum</i> .....	106
<i>Tabulka T: T–F (Q1–4)</i> .....	107
<i>Tabulka U: T–F predikce</i> .....	107
<i>Tabulka V: Po–Pt sum</i> .....	109
<i>Tabulka W: Po–Pt (Q1–4)</i> .....	109
<i>Tabulka X: Po–Pt predikce</i> .....	110
<i>Tabulka Y: kontexty chi-square (p-value)</i> .....	112
<i>Tabulka Z: kontexty (ne)potvrzující hypotézy</i> .....	113

### 8.1 Tabulka A: formální a aktuální členění

<b>formální členění</b>	<b>aktuální členění</b>
gramatická struktura	věcná souvislost
gramatický subjekt	východiště výpovědi, jádro
gramatický predikát	jádro výpovědi

## 8.2 Tabulka B: podmět, předmět

	podmět	předmět
<i>základní forma realizace</i>	elize	pronominalizace nebo plnovýznamové pojmenování
<i>elize</i>	ano	ne
<i>pronominalizace</i>	ano	ano
<i>zájmeno osobní on</i>	3. osoba	nedůrazové postavení předložkové pády
<i>zájmeno ukazovací ten</i>	ano	důrazové postavení nepředložkové pády
<i>plnovýznamové pojmenování</i>	ne	ne
<i>identifikátor ten</i>	ne	ano (identifikace) ne (jiné prostředky)

## 8.3 Tabulka C: „tematické“ posloupnosti v teorii centeringu

	$Cb(U_i) = Cb(U_{i-1})$ $Cb(U_{i-1}) = [?]$ <sup>89</sup>	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	continue	smooth-shift
$Cb(U_i) \neq Cp(U_i)$	retain	rough-shift

<sup>89</sup> Tento typ vztahu je využíván v případě první výpovědi kontextu, kdy není žádná předchozí výpověď, s níž by bylo provedeno srovnání.

**8.4 Tabulka D: Centering – analýza**

- a.
- Jeff pomáhal Dickovi mýt auto.*

Cp	[Jeff]
Cb	[?]
Cf	[Jeff, Dick, auto]
přechod	0

*Jeff myl okna, zatímco Dick voskoval.*

Cp	[Jeff]
Cb	[Jeff]
Cf	[Jeff, okna, Dick, auto]
přechod	continue

*Jeff vyleštil přední sklo.*

Cp	[Jeff]
Cb	[Jeff]
Cf	[Jeff, přední sklo]
přechod	continue

- b.
- Jeff pomáhal Dickovi mýt auto.*

Cp	[Jeff]
Cb	[?]
Cf	[Jeff, Dick, auto]
přechod	0

*Jeff myl okna, zatímco Dick voskoval.*

Cp	[Jeff]
Cb	[Jeff]
Cf	[Jeff, okna, Dick]
přechod	continue

*Dick taky leštil kapotu.*<sup>90</sup>

Cp	[Dick]
Cb	[Dick]
Cf	[Dick, kapota]
přechod	smooth-shift

c. *Tom si koupil nový počítač.*

Cp	[Tom]
Cb	[?]
Cf	[Tom, nový počítač]
přechod	0

*Tom počítač ukázal Johnovi.*

Cp	[Tom]
Cb	[Tom]
Cf	[Tom, počítač, John]
přechod	continue

*John Tomovi vysvětlil jeho vylepšené funkce.*

Cp	[John]
Cb	[Tom]
Cf	[John, Tom]
přechod	retain

## 8.5 Tabulka E: textová koreference (tfa)

	do t/c	do f	sum
z t/c	4 515	1 565	6 080
z f	31	15	46
sum	4 546	1 580	6 126

<sup>90</sup> Typ rough-shift by vypadal podobně, třetí věta v pořadí by zněla zhruba takto: *Dickem byla (vy)leštěna kapota., Kapota byla (vy)leštěna (Dickem).* V takovém případě by centra pozornosti byla rozložena takto: Cp [kapota], Cb [Dick], Cf [Dick, kapota], přechod: rough-shift. Tomuto typu přechodu se nevěnujeme, neboť se prakticky nevyskytuje. I v češtině je takové uspořádání diskurzu nepřirozené a nekompletní (informace o tom, že *Dickem byla vyleštěna kapota.* vyžaduje další pokračování výpovědi, např. *Dickem byla vyleštěna kapota, zatímco okna myl Jeff.*).

**8.6 Tabulka F: koreference v ef. br. (AČV)**

	do t/c	do f	sum
z t/c	1 178	853	2 031
z f	21	12	33
sum	1 199	865	2 064

**8.7 Tabulka G: koreference v ef. br. (VČ)**

	do Po	do Pt	jinam	sum
z Po	164	219	97	480
z Pt	261	203	114	578
odjinud	517	286	203	1 984
sum	942	708	414	2 064

**8.8 Tabulka H: TF + PoPt v ef. br.**

		do t/c 1 199			do f 865		
		661	333	205	281	375	209
		do Po	do Pt	jinam	do Po	do Pt	jinam
<b>z t/c</b>	<b>2031</b>	<b>653</b>	<b>325</b>	<b>200</b>	<b>280</b>	<b>369</b>	<b>204</b>
z Po	468	95	66	32	67	145	63
z Pt	571	201	114	54	58	87	57
odjinud	992	357	145	114	155	137	84
<b>z f</b>	<b>33</b>	<b>8</b>	<b>8</b>	<b>5</b>	<b>1</b>	<b>6</b>	<b>5</b>
z Po	12	2	5	1	0	3	1
z Pt	7	2	2	3	0	0	0
odjinud	14	4	1	1	1	3	4



## 8.9 Tabulka I: přehled výsledků měření (PDT)

	vzájemný poměr	X <sup>2</sup>	p-values
textová koreference			
z t/c	6 080:46	5 943.40	p < 2.2e-16
z f	132.17		
do t/c	4 546:1 580	1 436.04	p < 2.2e-16
do f	2.88		
ef. br.			
z t/c	2 031:33	1 934.11	p < 2.2e-16
z f	61.55		
do t/c	1 199: 865	54.05	p = 1.956e-13
do f	1.39		
z Po	480:578	9.08	p = 0.003
z Pt	0.83		
do Po	942:708	33.19	p = 8.377e-09
do Pt	1.33		

## 8.10 Tabulka J: schéma testovacích kontextů

aktuální členění	větněčlenská platnost
$[N1_{i1} - T, Po; N2_{i2} - F, Pt]_{V1} [(P_i)]_{V2}$ nastavení topikové	$[N1_{i1} - T, Po; N2_{i2} - F, Pt]_{V1} [(P_i)]_{V2}$ nastavení podmětové
$[N1_{i1} - F, Po; N2_{i2} - T, Pt]_{V1} [(P_i)]_{V2}$ nastavení fokusové	$[N1_{i1} - T, Pt; N2_{i2} - F, Po]_{V1} [(P_i)]_{V2}$ nastavení předmětové

## 8.11 Tabulka K: testovací kontext s notací

aktuální členění	větněčlenská platnost
$[N1_{i1} - T, Po; N2_{i2} - F, Pt]_{V1} [(P_i)]_{V2}$ Pan listonoš <sub>N1,T,Po</sub> dnes doručil objednaný časopis <sub>N2, F, Pt</sub> . Zpozdil se (on).	$[N1_{i1} - T, Po; N2_{i2} - F, Pt]_{V1} [(P_i)]_{V2}$ Pan listonoš <sub>N1,T,Po</sub> dnes doručil objednaný časopis <sub>N2, F, Pt</sub> . Zpozdil se (on).
$[N1_{i2} - F, Po; N2_{i1} - T, Pt]_{V1} [(P_i)]_{V2}$ Objednaný časopis <sub>N2,T,Pt</sub> dnes doručil pan listonoš <sub>N1, F, Po</sub> . Zpozdil se (on).	$[N1_{i2} - T, Pt; N2_{i1} - F, Po]_{V1} [(P_i)]_{V2}$ Panem listonošem <sub>N1,T,Pt</sub> byl dnes doručen objednaný časopis <sub>N2,F,Po</sub> . Zpozdil se (on).

**8.12 Tabulka L: sledované položky a zájmena**

	kontext	počet zájmen	typ zájmena	povrchově vyjádřené
1	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.	2	osobní osobní	N A
2	Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.	1	osobní	N
3	Pan listonoš dnes doručil objednaný časopis. Zpozdil se.	1	osobní	N
4	Liberální voliči podporují demokratické kandidáty. Bojují za občanské svobody, volný trh a zmírnění chudoby.	1	osobní	N
5	Palice odrazily útok seker. Jejich zástupci bojovali statečně.	1	přivlastňovací	A
6	Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země.	1	přivlastňovací	A
7	Paní doktorka netrpělivě očekává snaživou kolegyni. Její pacient není připraven na operaci.	1	přivlastňovací	A
8	Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.	1	osobní	N
9	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.	2	osobní osobní	N A
10	Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.	1	přivlastňovací	A
11	Petr řídí Volkswagen. Jezdí rychle.	1	osobní	N
12	Pittsburští Tučňáci předstihli newyorské Rangery. Jejich fanoušci byli překvapeni.	1	přivlastňovací	A
13	Policie vyšetřuje korupci. Její vliv ve veřejném prostoru je nezpochybnitelný.	1	přivlastňovací	A
14	Pořadatel nejprve uvítal na konferenci zahraničního profesora. Měl radost, že se potkali.	1	osobní	N
15	Národní knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	1	osobní	A
16	Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozířetelnost zachránila firmě reputaci.	1	přivlastňovací	A
17	Fotbalisté Bohemians pozvali na zápas kluky. Holky je obdivovaly.	1	osobní	A
18	Začínající autor na veletrhu představuje svůj první román. Velký zájem o něj nastrovala televizní reklama.	1	osobní	A
19	Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil ho.	2	osobní osobní	N A
20	Tomáš pozval na zápas Michala. Byl velký fanoušek fotbalu.	1	osobní	N

**8.13 Tabulka M: kódování výsledků (přijatelnost)**

kód	varianta lexikálního obsazení	varianta proměnné
A1	Knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	T, Po
A2	Fakultě darovala knihy knihovna. Bylo to pro ni výhodné.	F
A3	Knihovnou byly knihy darovány fakultě. Bylo to pro ni výhodné.	Pt

**8.14 Tabulka N: přijatelnost (akce)**

typ	kritérium	počet variant	varianty	počet kontextů	kontext	akce
I.	průměr i medián do 4.00	38	A1–D1, F1–M1, O1, P1; A2–P2; D3, H3–N3	7	D, H, I, J, K, L, M	zachovány jako testovací
II.	průměr i medián 4.00–4.50	3	B3, G3, P3	3	B, G, P	zachovány jako testovací
III.	průměr nebo medián 4.50–5.00, jediné z trojice	3	N1; F3, O3	3	F, N, O	zachovány jako testovací, upraveny po vzoru lépe hodnocených
IV.	více variant v trojici hodnoceno negativně (nad 4.50)	2	E1; E3	1	E	přesunuty do výplňkových kontextů, nahrazeny novým testovacím kontextem
V.	průměr i medián negativní (nad 5.00)	2	A3, C3	2	A, C	přesunuty do výplňkových kontextů, nahrazeny novým testovacím kontextem

**8.15 Tabulka O: upravené kontexty**

kontext	původní kontext	upravený kontext
F	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak jí koupila květiny.	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.
N	Palice odrazily útok seker. Jejich majitelé byli překvapeni.	Palice odrazily útok seker. Jejich zástupci bojovali statečně.
O	Paní primářka netrpělivě očekává vrchní sestru. Její dokumenty nejsou připraveny.	Paní doktorka netrpělivě očekává snaživou kolegyni. Její pacient není připraven na operaci.

**8.16 Tabulka P: nové kontexty**

kontext	původní kontext	nový kontext
A	Knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	Pan listonoš dnes doručil objednaný časopis. Zpozdil se.
C	Tomáš pozval na zápas Michala. Byl velký fanoušek.	Fotbalisté Bohemians pozvali na zápas kluky. Holky je obdivovaly.
E	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.	Pittsburští Tučňáci předstihli newyorské Rangery. Jejich fanoušci byli překvapeni.

**8.17 Tabulka Q: Q1–4**

	Q1 fialová	Q2 modrá	Q3 zelená	Q4 žlutá
T	kontexty 1–5	kontexty 6–10	kontexty 11–15	kontexty 6–10
F	kontexty 6–10	kontexty 1–5	kontexty 16–20	kontexty 11–15
Po	kontexty 11–15	kontexty 16–20	kontexty 1–5	kontexty 16–20
Pt	kontexty 16–20	kontexty 11–15	kontexty 6–10	kontexty 1–5

**8.18 Tabulka R: nezaznamenané hodnoty**

	aktuální členění	větný člen	
<i>T/F</i>	2 810	2 796	<i>Po/Pt</i>
<i>nezaznamenáno</i>	30	44	<i>nezaznamenáno</i>
<i>sum</i>	2 840	2 840	<i>sum</i>

**8.19 Tabulka S: T–F sum**

	T	F
sum	1 542	1 268
chi <sup>2</sup>	26.717	
p-value	2.355e-07	

**8.20 Tabulka T: T–F (Q1–4)**

	T	F	rozdíl	p-values	p < 0.05
Q1	321	385	64	0.016	A
Q2	439	255	184	2.858e-12	A
Q3	419	284	135	3.55e-07	A
Q4	363	344	19	0.475	N
	1 542	1 268	402	2.355e-07	A

**8.21 Tabulka U: T–F predikce**

	T > T	T > F	F > F	F > T
Q1	208	145	240	113
Q2	112	237	202	143
Q3	222	131	153	197
Q4	225	129	215	138

**8.22 Tabulka V: Po–Pt sum**

	Po	Pt
sum	1 701	1 095
$\chi^2$	131.343	
p-value	2.2e-16	

**8.23 Tabulka W: Po–Pt (Q1–4)**

	Po	Pt	rozdíl	p-values	p < 0.05
Q1	495	204	291	2.2e-16	A
Q2	464	228	236	2.2e-16	A
Q3	406	298	108	4.693e-05	A
Q4	336	365	29	0.273	N
	1 701	1 095	606	2.2e-16	A

**8.24 Tabulka X: Po–Pt predikce**

	Po > Po	Po > Pt	Pt > Pt	Pt > Po
Q1	224	129	75	271
Q2	236	109	119	228
Q3	212	141	157	194
Q4	100	251	114	236

**8.25 Tabulka Y: chi-square (p-values)<sup>91</sup>**

kontexty	T	F	T–F	Po	Pt	Po–Pt
1	65 5	14 55	p < 0.05	62 8	20 49	p < 0.05
2	60 11	10 58	p < 0.05	56 15	45 25	p = 0.083
3	25 46	32 38	p = 0.272	32 39	61 9	p < 0.05
4	21 49	39 28	p < 0.05	25 45	58 12	p < 0.05
5	37 34	48 23	p = 0.87	37 34	52 19	p < 0.05
6	68 3	2 69	p < 0.05	64 3	10 59	p < 0.05
7	40 31	34 35	p = 0.504	46 23	44 27	p = 0.687
8	65 6	10 61	p < 0.05	62 8	39 32	p < 0.05
9	45 26	15 56	p < 0.05	47 22	54 24	p = 1.000
10	7 63	52 19	p < 0.05	17 53	56 15	p < 0.05
11	50 21	13 57	p < 0.05	54 17	57 13	p = 0.566
12	45 26	40 30	p = 0.559	45 25	40 28	p = 0.628
13	16 53	54 17	p < 0.05	12 58	60 10	p < 0.05
14	66 5	2 69	p < 0.05	65 6	17 54	p < 0.05
15	45 26	29 42	p < 0.05	48 23	54 14	p = 0.167
16	43 27	25 44	p < 0.05	39 31	53 13	p < 0.05
17	23 48	37 34	p < 0.05	21 49	41 28	p < 0.05
18	14 54	56 14	p < 0.05	10 60	52 17	p < 0.05
19	10 59	62 8	p < 0.05	10 60	64 7	p < 0.05
20	22 49	17 53	p = 0.483	20 51	61 10	p < 0.05

<sup>91</sup> Pořadí kontextů odpovídá pořadí uvedenému v příloze Sledované položky (10.4). Označení: podtrženo = signifikantní v rámci obou proměnných; *kurzíva* = nesignifikantní v rámci obou proměnných.

**8.26 Tabulka Z: kontexty (ne)potvrzující hypotézy**

skupina	počet	kontexty
I - potvrzující hypotézu	7	1, 2, 6, 7, 8, 12, 14
II - nepotvrzující hypotézu	8	3, 4, 5, 10, 13, 17, 18, 19
III - AČV potvrzující, VČ nepotvrzující hypotézu	5	9, 11, 15, 16, 20

## 9 Seznam příkladů

- a. *[Galileo]T veřejně podporoval [Koperníka]F. Věřil (on), že Země se otáčí kolem Slunce. [Koperníka]T veřejně podporoval [Galileo]F. Věřil, že Země se otáčí kolem Slunce.*  
*[Galileo]Po veřejně podporoval [Koperníka]Pt. Věřil (on), že Země se otáčí kolem Slunce.*  
*[Galileem]Pt byl veřejně podporován [Koperník]Po. Věřil (on), že Země se otáčí kolem Slunce.*
- b. *Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila (ona) schůzku.*
- c. *Petra musela oloupat jablka. Byla (ona) otrávená.*
- d. *Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil (on) ho.*
- e. *Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.*
- f. *Pittsburští Tučňáci odrazili newyorské Rangery. Jejich fanoušci byli překvapeni.*
- g. *Národní knihovna darovala knihy fakultě. Bylo to pro ni výhodné.*
- h. *[Karolína]<sub>i1</sub> chtěla osobně navštívit [babičku]<sub>i2</sub>. [Pro]<sub>i1/i2</sub> odjela do Prahy.*
- i. *?[Karolína]<sub>i1</sub> chtěla osobně navštívit [babičku]<sub>i2</sub>. [Ona]<sub>i2</sub> odjela do Prahy.<sup>92</sup>*
- j. *[Karolína]<sub>i1</sub> chtěla osobně navštívit [babičku]<sub>i2</sub>. [Ona]<sub>i2</sub> ale odjela do Prahy.*
- k. *[Karolína]<sub>i1</sub> chtěla osobně navštívit [babičku]<sub>i2</sub>. [Ta]<sub>i2</sub> ale odjela do Prahy.*
- l. *Hana šla převzít službu od [primáře Vojtěcha]<sub>i</sub>. V kanceláři však zjistila, že on<sub>i</sub> bude mít službu až zítra.*
- m. *Hana<sub>i</sub> šla převzít službu od primáře Vojtěcha. V kanceláři však zjistila, že ona<sub>i</sub> bude mít službu až zítra.*
- n. *Hana<sub>i1</sub> šla převzít službu od [primáře Vojtěcha]<sub>i2</sub>. V kanceláři však zjistila, že [PRO]<sub>i1/i2</sub> bude mít službu až zítra.*
- o. *TOP 09 je nyní v Poslanecké sněmovně. O její budoucnost se neobávám.*

<sup>92</sup> Vzhledem k tomu, že kontext B bych osobně považovala za negramatický, nejsem schopna posoudit, nakolik je koreferenční řetězec jednoznačný. Domnívám se ale, že možnost odkazování na [Karolína]<sub>i1</sub> je velice oslabena, pokud ne zcela vyloučena.



- p. Češi chtějí vidět Jaromíra Jágra. Jejich touha vyprodala O2 arénu.*
- q. Norové odmítli Evropskou unii. Ted' (oni) jí dávají Nobelovu cenu míru.*
- r. Do České filharmonie se vrací bývalý dirigent. Jeho dva koncerty vyvolaly kladné ohlasy.*
- s. Budova nového pavilonu v botanické zahradě bude patřit různým. Ty jsou novým logem zahrady.*
- t. České fotbalisty čeká další bitva. Snad (ona) dopadne dobře.*
- u. (Roger Federer odehrál exhibiční zápas s Rafaelem Nadalem. Publikum mu tleskalo ve stoje.) Komu tleskalo publikum vestoje?*
- v. (Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.) Kdo si domluvil schůzku s kým?*
- w. (Vize jednotné Ameriky byla překryta novým přístupem. Apeluje na třídní nevraživost.) Co apeluje na třídní nevraživost?*
- x. (Volební výsledky mají moc zamíchat dozorčí radou. Přímo či nepřímá závisí na hlasech voličů.) Kdo/Co závisí na hlasech voličů?*
- y. (Autosalon propaguje nový vůz. Jeho úspěšnost závisí na reklamě.) Čí úspěšnost závisí na reklamě?*
- z. (Pan listonoš dnes doručil objednaný časopis. Zpozdil se.) Kdo/Co se zpozdil/o?*
- aa. (Postřelenou dívku převezli do jiné nemocnice. Je lépe hlídána.) Kdo/Co je lépe hlídáný/é?*

## 10 Přílohy

<i>10.1 Algoritmus a příklad (BFP 1987: 159) .....</i>	<i>147</i>
<i>10.2 Kontexty testování přijatelnosti .....</i>	<i>148</i>
<i>10.3 Výsledky testování přijatelnosti .....</i>	<i>150</i>
<i>10.4 Sledované položky.....</i>	<i>158</i>
<i>10.5 T-test.....</i>	<i>161</i>
<i>10.6 Dotazníkové položky .....</i>	<i>162</i>
<i>10.7 Otázky a možnosti odpovědi .....</i>	<i>167</i>
<i>10.8 T-F podle nastavení kontextu (chi-square, p-values).....</i>	<i>168</i>
<i>10.9 Po-Pt podle nastavení kontextu (chi-square, p-values).....</i>	<i>168</i>

## 10.1 Algoritmus a příklad (BFP 1987: 159)

1. **CONSTRUCT THE PROPOSED ANCHORS** for  $U_n$  **EXAMPLE:** She doesn't believe him. ( $U_{n+4}$  from figure 2)
- (a) Create set of referring expressions (RE's).  $\Rightarrow ([A4] [A5])$
  - (b) Order RE's by grammatical relation.  $\Rightarrow ([A4] [A5])$
  - (c) Create set of possible forward center ( $Cf$ ) lists. Expand each element of (b) according to whether it is a pronoun or a proper name. Expand pronouns into set with entry for each discourse entity which matches its agreement features and expand proper nouns into a set with an entry for each possible referent. These expansions are a way of encoding a disjunction of possibilities.  $\Rightarrow ([FRIEDMAN:A4] [POLLARD:A5])$
  - (d) Create list of possible backward centers ( $Cb$ 's). This is taken as the entities from  $Cf(U_{n-1})$  plus an additional entry of NIL to allow the possibility that we will not find a  $Cb$  for the current utterance.  $\Rightarrow ([POLLARD:A2] [FRIEDMAN:A3] [RAISE:X1] NIL)$
  - (e) Create the proposed anchors. ( $Cb$ - $Cf$  combinations from the cross-product of the previous two steps)  $\Rightarrow$  There are four possible  $\langle Cb, Cf \rangle$  pairs for this utterance.
    - i.  $\langle [POLLARD:A2], ([FRIEDMAN:A4] [POLLARD:A5]) \rangle$
    - ii.  $\langle [FRIEDMAN:A3], ([FRIEDMAN:A4] [POLLARD:A5]) \rangle$
    - iii.  $\langle [RAISE:X1], ([FRIEDMAN:A4] [POLLARD:A5]) \rangle$
    - iv.  $\langle NIL, ([FRIEDMAN:A4] [POLLARD:A5]) \rangle$
2. **FILTER THE PROPOSED ANCHORS**
- For each anchor in our list of proposed anchors we apply the following three filters. If it passes each filter then it is still a possible anchor for the current utterance.
- (a) Filter by contraindices. That is, if we have proposed the same antecedent for two contraindiced pronouns or if we have proposed an antecedent for a pronoun which it is contraindiced with, eliminate this anchor from consideration.  $\Rightarrow$  This filter doesn't eliminate any of the proposed anchors in this example. Even though [A4] and [A5] are contraindiced we have not proposed the same co-specifier due to agreement.
  - (b) Go through  $Cf(U_{n-1})$  keeping (in order) those which appear in the proposed  $Cf$  list of the anchor. If the proposed  $Cb$  of the anchor does not equal the first element of this constructed list then eliminate this anchor. This guarantees that the  $Cb$  will be the highest ranked element of the  $Cf(U_{n-1})$  realized in the current utterance. (This corresponds to constraint 3 given in section 1)  $\Rightarrow$  This filter eliminates proposed anchors ii, iii, iv.
  - (c) If none of the entities realized as pronouns in the proposed  $Cf$  list equals the proposed  $Cb$  then eliminate this anchor. This guarantees that if any element is realized as a pronoun then the  $Cb$  is realized as a pronoun. (If there are no pronouns in the proposed  $Cf$  list then the anchor passes this filter. This corresponds to rule 1 in section 1). This rule could be implemented as a preference strategy rather than a strict filter.  $\Rightarrow$  This filter doesn't eliminate any of the proposed anchors. The proposed  $Cb$  was realized as a pronoun.
3. **CLASSIFY and RANK**
- (a) Classify each anchor on the list of proposed anchors by the transitions as described in section 1 taking  $U_{n-1}$  to be the previous utterance and  $U_n$  to be the one we are currently working on.  $\Rightarrow$  Anchor i is classified as a retention based on the transition state definition.
  - (b) Rank each proposed anchor using the extended ranking in section 2. Set  $Cb(U_n)$  to the proposed  $Cb$  and  $Cf(U_n)$  to proposed  $Cf$  of the most highly ranked anchor.  $\Rightarrow$  Anchor i is the most highly ranked anchor (trivially).

Figure 7 : Algorithm and Example

**10.2 Kontexty testování přijatelnosti**

kód	položka
A1	Knihovna darovala knihy fakultě. Bylo to pro ni výhodné.
B1	Pořadatel nejprve uvítal na konferenci zahraničního profesora. Měl radost, že se potkali.
C1	Tomáš pozval na zápas Michala. Byl velký fanoušek fotbalu.
D1	Petr řídí Volkswagen. Jezdí rychle.
E1	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.
F1	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak jí koupila květiny.
G1	Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.
H1	Liberální voliči podporují demokratické kandidáty. Bojují za občanské svobody, volný trh a zmírnění chudoby.
I1	Policie vyšetřuje korupci. Její vliv ve veřejném prostoru je nezpochybnitelný.
J1	Začínající autor na veletrhu představuje svůj první román. Velký zájem o něj nastartovala televizní reklama.
K1	Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Velice rychle ho proslavil.
L1	Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.
M1	Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.
N1	Palice odrazily útok seker. Jejich majitelé byli překvapeni.
O1	Paní primářka netrpělivě očekává vrchní sestru. Její dokumenty nejsou připraveny.
P1	Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.
A2	Fakultě darovala knihy knihovna. Bylo to pro ni výhodné.
B2	Zahraničního profesora nejprve uvítal na konferenci pořadatel. Měl radost, že se potkali.
C2	Michala pozval na zápas Tomáš. Byl velký fanoušek fotbalu.
D2	Volkswagen řídí Petr. Jezdí rychle.
E2	Karla srdečně přivítal pan domácí. Půjčil mu peníze.
F2	Kláru na volné místo doporučila dlouholetá kolegyně. Pak jí koupila květiny.
G2	Koperníka veřejně podporoval Galileo. Věřil, že Země se otáčí kolem Slunce.
H2	Demokratické kandidáty podporují liberální voliči. Bojují za občanské svobody, volný trh a zmírnění chudoby.
I2	Korupci vyšetřuje policie. Její vliv ve veřejném prostoru je nezpochybnitelný.
J2	Svůj první román představuje na veletrhu začínající autor. Velký zájem o něj nastartovala televizní reklama.
K2	Moderní dům ve Francouzské čtvrti navrhl známý architekt Gaudí. Velice rychle ho proslavil.
L2	Dvacet nových zaměstnanců přijali minulý měsíc do naší firmy personalisté. Jejich pracovitost je nepřekonatelná.
M2	Nového generálního ředitele Alexandra Schramma jmenoval letos Ronald McDonald. Jeho prozřetelnost zachránila firmě reputaci.
N2	Útok seker odrazily palice. Jejich majitelé byli překvapeni.
O2	Vrchní sestru netrpělivě očekává paní primářka. Její dokumenty nejsou připraveny.
P2	Aničku pochválila paní učitelka. Měla radost, že se jí cvičení povedlo.
A3	Knihovnou byly knihy darovány fakultě. Bylo to pro ni výhodné.
B3	Pořadatelem byl nejprve na konferenci uvítán zahraniční profesor. Měl radost, že se

	potkali.
C3	Tomášem byl na zápas pozván Michal. Byl velký fanoušek fotbalu.
D3	Petrovi se dobře řídí Volkswagen. Jezdí rychle.
E3	Panem domácím byl srdečně přivítán Karel. Půjčil mu peníze.
F3	Dlouholetou kolegyní byla na volné místo doporučena Klára. Pak jí koupila květiny.
G3	Galileem byl veřejně podporován Koperník. Věřil, že Země se otáčí kolem Slunce.
H3	Liberálními voliči jsou podporováni demokratičtí kandidáti. Bojují za občanské svobody, volný trh a zmírnění chudoby.
I3	Policií je vyšetřována korupce. Její vliv ve veřejném prostoru je nezpochybnitelný.
J3	Začínajícím autorem je na veletrhu představován jeho první román. Velký zájem o něj nastartovala televizní reklama.
K3	Známým architektem Gaudím byl navržen moderní dům ve Francouzské čtvrti. Velice rychle ho proslavil.
L3	Personalisty bylo minulý měsíc do naší firmy přijato dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.
M3	Ronaldem McDonaldeem byl letos jmenován nový generální ředitel Alexandr Schramm. Jeho prozřetelnost zachránila firmě reputaci.
N3	Palicemi byl odražen útok seker. Jejich majitelé byli překvapeni.
O3	Paní primářkou je netrpělivě očekávána vrchní sestra. Její dokumenty nejsou připraveny.
P3	Paní učitelkou byla pochválena Anička. Měla radost, že se jí cvičení povedlo.

## 10.3 Výsledky testování přijatelnosti

	PŘIJATELNOST KONTEXTŮ	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	průměr	median
		Pohlaví																					
		Věk																					
A1	Knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	2	3	3	6	2	1	5	7	1	4	2	1	3	5	1	1	2	3	3	1	2,80	2,50
B1	Pořadatel nejprve uvítal na konferenci zahraničního profesora. Měl radost, že se potkali.	2	6	2	1	2	3	4	1	3	2	4	1	4	4	2	1	1	3	4	1	2,55	2,00
C1	Tomáš pozval na zápas Michala. Byl velký fanoušek fotbalu.	3	4	5	1	5	3	5	7	1	6	2	1	2	7	2	1	2	4	5	5	3,55	3,50
D1	Petr řídí Volkswagen. Jezdí rychle.	3	5	2	1	5	6	2	7	3	7	3	1	2	4	1	1	1	2	2	6	3,20	2,50
E1	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.	6	6	4	1	5	5	6	7	3	6	5	1	3	6	2	1	2	4	6	6	4,25	5,00
F1	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak jí koupila květiny.	7	5	4	1	3	2	6	1	7	4	3	4	4	7	2	1	2	1	1	6	3,55	3,50
G1	Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.	1	2	2	1	1	1	1	1	1	6	2	1	4	7	1	1	2	1	5	2	2,15	1,00

H1	Liberální voliči podporují demokratické kandidáty. Bojují za občanské svobody, volný trh a zmírnění chudoby.	1	3	2	1	2	2	1	1	5	2	4	1	3	7	1	1	1	3	1	6	2,40	2,00
I1	Policie vyšetřuje korupci. Její vliv ve veřejném prostoru je nezpochybnitelný.	4	2	3	1	3	6	5	7	7	6	6	1	3	7	2	1	1	2	1	6	3,70	3,00
J1	Začínající autor na veletrhu představuje svůj první román. Velký zájem o něj nastartovala televizní reklama.	5	1	2	2	5	5	2	7	3	2	4	1	2	5	2	2	2	3	1	2	2,90	2,00
K1	Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Velice rychle ho proslavil.	7	7	3	1	2	7	2	7	7	6	4	3	5	6	2	1	1	5	1	2	3,95	3,50
L1	Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.	4	7	3	1	3	5	2	7	1	6	6	1	4	5	1	3	2	4	2	4	3,55	3,50

M 1	Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.	2	3	2	1	3	7	1	7	1	5	6	1	2	3	2	1	1	1	4	5	2,90	2,00
N1	Palice odrazily útok seker. Jejich majitelé byli překvapeni.	7	7	2	1	6	6	4	7	7	2	4	7	7	7	2	1	1	2	6	2	4,40	5,00
O1	Paní primářka netrpělivě očekává vrchní sestru. Její dokumenty nejsou připraveny.	1	1	2	1	4	7	1	7	7	7	3	1	2	7	2	1	1	3	1	2	3,05	2,00
P1	Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.	2	1	2	1	1	4	1	7	1	2	2	1	1	7	1	1	2	2	1	2	2,10	1,50
A2	Fakultě darovala knihy knihovna. Bylo to pro ni výhodné.	6	2	2	5	1	2	2	1	7	4	6	4	3	2	2	3	3	3	5	2	3,25	3,00
B2	Zahraničního profesora nejprve uvítal na konferenci pořadatel. Měl radost, že se potkali.	6	4	2	4	2	1	5	1	7	4	5	1	1	2	2	1	2	1	5	2	2,90	2,00
C2	Michala pozval na zápas Tomáš. Byl velký fanoušek fotbalu.	6	4	2	3	1	3	2	1	7	4	2	7	4	4	2	1	1	1	2	6	3,15	2,50
D2	Volkswagen řídí Petr. Jezdí rychle.	7	1	4	2	4	4	2	1	7	7	4	7	6	4	2	1	1	1	1	6	3,60	4,00



E2	Karla srdečně přivítal pan domácí. Půjčil mu peníze.	7	4	4	2	2	4	2	1	7	7	5	1	2	6	2	1	1	1	3	6	3,40	2,50
F2	Kláru na volné místo doporučila dlouholetá kolegyně. Pak jí koupila květiny.	3	2	4	1	1	1	1	1	7	3	4	3	3	5	2	1	1	1	1	2	2,35	2,00
G2	Koperník veřejně podporoval Galileo. Věřil, že Země se otáčí kolem Slunce.	1	3	5	1	1	1	2	1	1	6	1	3	2	3	2	2	1	1	3	4	2,20	2,00
H2	Demokratické kandidáty podporují liberální voliči. Bojují za občanské svobody, volný trh a zmírnění chudoby.	1	3	5	3	3	1	4	1	7	2	3	7	2	4	2	1	1	1	5	2	2,90	2,50
I2	Korupci vyšetřuje policie. Její vliv ve veřejném prostoru je nezpochybnitelný.	1	2	3	2	5	2	4	1	7	6	4	1	4	5	2	1	1	1	6	6	3,20	2,50
J2	Svůj první román představuje na veletrhu začínající autor. Velký zájem o něj nastartovala televizní reklama.	3	3	3	5	1	1	6	1	7	3	4	7	2	6	1	1	1	1	3	2	3,05	3,00
K2	Moderní dům ve Francouzské čtvrti navrhl známý architekt Gaudí. Velice rychle ho proslavil.	3	1	5	1	2	3	1	7	7	2	5	1	3	3	2	1	1	1	5	2	2,80	2,00

L2	Dvacet nových zaměstnanců přijali minulý měsíc do naší firmy personalisté. Jejich pracovitost je nepřekonatelná.	6	2	5	4	3	3	3	1	1	6	5	2	3	6	2	2	2	1	6	3	3,30	3,00
M 2	Nového generálního ředitele Alexandra Schramma jmenoval letos Ronald McDonald. Jeho prozřetelnost zachránila firmě reputaci.	1	2	4	1	2	5	2	1	7	7	5	7	2	3	2	1	2	1	1	5	3,05	2,00
N2	Útok seker odrazily palice. Jejich majitelé byli překvapeni.	6	6	3	1	7	3	2	1	7	4	5	1	4	7	2	1	1	1	1	2	3,25	2,50
O2	Vrchní sestru netrpělivě očekává paní primářka. Její dokumenty nejsou připraveny.	1	2	4	2	3	1	2	7	7	7	2	7	1	7	2	1	1	1	5	2	3,25	2,00
P2	Aničku pochválila paní učitelka. Měla radost, že se jí cvičení povedlo.	4	1	5	5	2	1	2	1	1	1	2	1	1	6	1	2	2	1	1	2	2,10	1,50
A3	Knihovnou byly knihy darovány fakultě. Bylo to pro ni výhodné.	7	7	5	7	4	6	6	7	7	6	3	4	7	7	3	3	2	2	7	2	5,10	6,00
B3	Pořadatelem byl nejprve na konferenci uvítán zahraniční profesor. Měl radost, že se potkali.	2	6	5	6	3	7	5	7	7	6	4	3	7	4	2	1	1	1	5	2	4,20	4,50

C3	Tomášem byl na zápas pozván Michal. Byl velký fanoušek fotbalu.	6	5	6	7	6	6	5	7	7	6	2	7	7	6	3	6	3	4	7	2	5,40	6,00
D3	Petrovi se dobře řídí Volkswagen. Jezdí rychle.	1	4	2	4	4	6	1	1	7	6	3	1	2	4	2	1	1	1	2	6	2,95	2,00
E3	Panem domácím byl srdečně přivítán Karel. Půjčil mu peníze.	3	6	5	7	5	7	5	7	7	7	2	1	6	7	2	4	1	4	7	2	4,75	5,00
F3	Dlouholetou kolegyní byla na volné místo doporučena Klára. Pak jí koupila květiny.	1	5	6	7	6	5	5	7	7	6	3	7	7	7	2	4	1	2	3	3	4,70	5,00
G3	Galileem byl veřejně podporován Koperník. Věřil, že Země se otáčí kolem Slunce.	2	6	3	7	4	6	6	1	7	6	4	7	6	5	2	4	2	5	4	2	4,45	4,50
H3	Liberálními voliči jsou podporováni demokratičtí kandidáti. Bojují za občanské svobody, volný trh a zmírnění chudoby.	1	3	3	7	3	4	6	1	7	1	4	7	6	2	2	3	2	4	1	2	3,45	3,00
I3	Policií je vyšetřována korupce. Její vliv ve veřejném prostoru je nezpochybnitelný.	2	3	5	7	2	2	6	1	7	6	5	3	4	6	3	1	2	4	2	3	3,70	3,00

J3	Začínajícím autorem je na veletrhu představován jeho první román. Velký zájem o něj nastrovala televizní reklama.	2	4	4	6	3	5	4	1	7	2	3	7	3	2	2	5	1	5	1	2	3,45	3,00
K3	Známým architektem Gaudím byl navržen moderní dům ve Francouzské čtvrti. Velice rychle ho proslavil.	2	2	4	6	2	3	5	1	7	6	3	2	2	2	2	4	1	3	4	3	3,20	3,00
L3	Personalisty bylo minulý měsíc do naší firmy přijato dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.	2	6	3	6	3	2	3	1	1	3	4	1	2	3	1	3	1	2	2	4	2,65	2,50
M3	Ronaldem McDonalodem byl letos jmenován nový generální ředitel Alexandr Schramm. Jeho prozířetelnost zachránila firmě reputaci.	2	3	4	7	3	2	4	1	1	4	4	1	4	3	2	3	1	3	3	2	2,85	3,00
N3	Palicemi byl odražen útok seker. Jejich majitelé byli překvapeni.	4	1	5	7	6	3	1	1	7	2	5	7	2	7	2	2	1	2	2	4	3,55	2,50
O3	Paní primářkou je netrpělivě očekávána vrchní sestra. Její dokumenty nejsou připraveny.	7	1	6	7	5	2	5	7	7	7	3	7	6	6	3	1	1	5	5	5	4,80	5,00

P3	Paní učitelkou byla pochválena Anička. Měla radost, že se jí cvičení povedlo.	7	1	5	7	7	7	6	1	7	3	4	3	6	4	3	2	2	5	7	2	4,45	4,50
----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	------	------

**10.4 Sledované položky**

	AKTUÁLNÍ ČLENĚNÍ		VĚTNÉ ČLENY	
	TOPIK	FOKUS	PODMĚT	PŘEDMĚT
1	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.	Kláru na volné místo doporučila dlouholetá kolegyně. Pak si s ní domluvila schůzku.	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.	Dlouholetou kolegyní byla na volné místo doporučena Klára. Pak si s ní domluvila schůzku.
2	Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.	Koperníka veřejně podporoval Galileo. Věřil, že Země se otáčí kolem Slunce.	Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.	Galileem byl veřejně podporován Koperník. Věřil, že Země se otáčí kolem Slunce.
3	Pan listonoš dnes doručil objednaný časopis. Zpozdil se.	Objednaný časopis dnes doručil pan listonoš. Zpozdil se.	Pan listonoš dnes doručil objednaný časopis. Zpozdil se.	Panem listonošem byl dnes doručen objednaný časopis. Zpozdil se.
4	Liberální voliči podporují demokratické kandidáty. Bojují za občanské svobody, volný trh a zmírnění chudoby.	Demokratické kandidáty podporují liberální voliči. Bojují za občanské svobody, volný trh a zmírnění chudoby.	Liberální voliči podporují demokratické kandidáty. Bojují za občanské svobody, volný trh a zmírnění chudoby.	Liberálními voliči jsou podporováni demokratičtí kandidáti. Bojují za občanské svobody, volný trh a zmírnění chudoby.
5	Palice odrazily útok seker. Jejich zástupci bojovali statečně.	Útok seker odrazily palice. Jejich zástupci bojovali statečně.	Palice odrazily útok seker. Jejich zástupci bojovali statečně.	Palicemi byl odražen útok seker. Jejich zástupci bojovali statečně.
6	Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země.	Prezidenta Václava Klause po deseti letech střídá Miloš Zeman. Jeho příznivci se na inauguraci sjíždějí z celé země.	Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země.	Milošem Zemanem je po deseti letech střídán prezident Václav Klaus. Jeho příznivci se na inauguraci sjíždějí z celé země.
7	Paní doktorka netrpělivě očekává snaživou kolegyni. Její pacient není připraven na operaci.	Snaživou kolegyni netrpělivě očekává paní doktorka. Její pacient není připraven na operaci.	Paní doktorka netrpělivě očekává snaživou kolegyni. Její pacient není připraven na operaci.	Paní doktorkou je netrpělivě očekávána snaživá kolegyně. Její pacient není připraven na operaci.
8	Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.	Aničku pochválila paní učitelka. Měla radost, že se jí cvičení povedlo.	Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.	Paní učitelkou byla pochválena Anička. Měla radost, že se jí cvičení povedlo.
9	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.	Karla srdečně přivítal pan domácí. Půjčil mu peníze.	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.	Panem domácím byl srdečně přivítán Karel. Půjčil mu peníze.

10	Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.	Dvacet nových zaměstnanců přijali minulý měsíc do naší firmy personalisté. Jejich pracovitost je nepřekonatelná.	Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.	Personalisty bylo minulý měsíc do naší firmy přijato dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.
11	Petr řídí Volkswagen. Jezdí rychle.	Volkswagen řídí Petr. Jezdí rychle.	Petr řídí Volkswagen. Jezdí rychle.	Petrovi se dobře řídí Volkswagen. Jezdí rychle.
12	Pittsburští Tučňáci předstihli newyorské Rangery. Jejich fanoušci byli překvapeni.	Newyorské Rangery předstihli pittsburští Tučňáci. Jejich fanoušci byli překvapeni.	Pittsburští Tučňáci předstihli newyorské Rangery. Jejich fanoušci byli překvapeni.	Pittsburskými Tučňáky byli předstiženi newyorští Rangeři. Jejich fanoušci byli překvapeni.
13	Policie vyšetřuje korupci. Její vliv ve veřejném prostoru je nezpochybnitelný.	Korupci vyšetřuje policie. Její vliv ve veřejném prostoru je nezpochybnitelný.	Policie vyšetřuje korupci. Její vliv ve veřejném prostoru je nezpochybnitelný.	Policií je vyšetřována korupce. Její vliv ve veřejném prostoru je nezpochybnitelný.
14	Pořadatel nejprve uvítal na konferenci zahraničního profesora. Měl radost, že se potkali.	Zahraničního profesora nejprve uvítal na konferenci pořadatel. Měl radost, že se potkali.	Pořadatel nejprve uvítal na konferenci zahraničního profesora. Měl radost, že se potkali.	Pořadatelem byl nejprve na konferenci uvítán zahraniční profesor. Měl radost, že se potkali.
15	Národní knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	Fakultě darovala knihy Národní knihovna. Bylo to pro ni výhodné.	Národní knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	Národní knihovnou byla obdarována fakulta. Bylo to pro ni výhodné.
16	Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.	Nového generálního ředitele Alexandra Schramma jmenoval letos Ronald McDonald. Jeho prozřetelnost zachránila firmě reputaci.	Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.	Ronaldem McDonaledem byl letos jmenován nový generální ředitel Alexandr Schramm. Jeho prozřetelnost zachránila firmě reputaci.
17	Fotbalisté Bohemians pozvali na zápas kluky. Holky je obdivovaly.	Kluky pozvali na zápas fotbalisté Bohemians. Holky je obdivovaly.	Fotbalisté Bohemians pozvali na zápas kluky. Holky je obdivovaly.	Fotbalisty Bohemians byli na zápas pozváni kluci. Holky je obdivovaly.
18	Začínající autor na veletrhu představuje svůj první román. Velký zájem o něj nastartovala televizní reklama.	Svůj první román představuje na veletrhu začínající autor. Velký zájem o něj nastartovala televizní reklama.	Začínající autor na veletrhu představuje svůj první román. Velký zájem o něj nastartovala televizní reklama.	Začínajícím autorem je na veletrhu představován jeho první román. Velký zájem o něj nastartovala televizní reklama.

19	Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil ho.	Moderní dům ve Francouzské čtvrti navrhl známý architekt Gaudí. Proslavil ho.	Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil ho.	Známým architektem Gaudím byl navržen moderní dům ve Francouzské čtvrti. Proslavil ho.
20	Tomáš pozval na zápas Michala. Byl velký fanoušek fotbalu.	Michala pozval na zápas Tomáš. Byl velký fanoušek fotbalu.	Tomáš pozval na zápas Michala. Byl velký fanoušek fotbalu.	Tomášem byl na zápas pozván Michal. Byl velký fanoušek fotbalu.



**10.5 T-test**

<b>p-values</b>	T–F	Po–Pt
A	0.426	0.000
B	0.517	0.008
C	0.536	0.003
D	0.576	0.704
E	0.212	0.459
F	0.059	0.102
G	0.925	0.000
H	0.403	0.101
I	0.477	1.000
J	0.805	0.339
K	0.107	0.262
L	0.678	0.115
M	0.825	0.930
N	0.140	0.269
O	0.799	0.023
P	1.000	0.001

**10.6 Dotazníkové položky**

		<b>Quiz 1</b>	<b>Quiz 2</b>	<b>Quiz 3</b>	<b>Quiz 4</b>
1	V	Obraz Emila Filly pokořil rekord. Získal ho nový majitel.	Obraz Emila Filly pokořil rekord. Získal ho nový majitel.	Obraz Emila Filly pokořil rekord. Získal ho nový majitel.	Obraz Emila Filly pokořil rekord. Získal ho nový majitel.
2	V	Roger Federer odehrál exhibiční zápas s Rafaellem Nadalem. Publikum mu tleskalo ve stoje.	Roger Federer odehrál exhibiční zápas s Rafaellem Nadalem. Publikum mu tleskalo ve stoje.	Roger Federer odehrál exhibiční zápas s Rafaellem Nadalem. Publikum mu tleskalo ve stoje.	Roger Federer odehrál exhibiční zápas s Rafaellem Nadalem. Publikum mu tleskalo ve stoje.
3	V	Vize jednotné Ameriky byla překryta novým přístupem. Apeluje na třídní nevraživost.	Vize jednotné Ameriky byla překryta novým přístupem. Apeluje na třídní nevraživost.	Vize jednotné Ameriky byla překryta novým přístupem. Apeluje na třídní nevraživost.	Vize jednotné Ameriky byla překryta novým přístupem. Apeluje na třídní nevraživost.
4	V	Volební výsledky mají moc zamíchat dozorčí radou. Přímo či nepřímo závisí na hlasech voličů.	Volební výsledky mají moc zamíchat dozorčí radou. Přímo či nepřímo závisí na hlasech voličů.	Volební výsledky mají moc zamíchat dozorčí radou. Přímo či nepřímo závisí na hlasech voličů.	Volební výsledky mají moc zamíchat dozorčí radou. Přímo či nepřímo závisí na hlasech voličů.
5	V	Barley nabídl whiskey bratrovi. Chutnala mu.	Barley nabídl whiskey bratrovi. Chutnala mu.	Barley nabídl whiskey bratrovi. Chutnala mu.	Barley nabídl whiskey bratrovi. Chutnala mu.
6	S	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.	Kláru na volné místo doporučila dlouholetá kolegyně. Pak si s ní domluvila schůzku.	Dlouholetá kolegyně doporučila na volné místo Kláru. Pak si s ní domluvila schůzku.	Dlouholetou kolegyní byla na volné místo doporučena Klára. Pak si s ní domluvila schůzku.
7	V	Moderátor útočil na starostu. Nechtěl odpovídat na otázky.	Moderátor útočil na starostu. Nechtěl odpovídat na otázky.	Moderátor útočil na starostu. Nechtěl odpovídat na otázky.	Moderátor útočil na starostu. Nechtěl odpovídat na otázky.
8	V	Šéf sledoval každý krok svého asistenta. Nedůvěřoval mu.	Šéf sledoval každý krok svého asistenta. Nedůvěřoval mu.	Šéf sledoval každý krok svého asistenta. Nedůvěřoval mu.	Šéf sledoval každý krok svého asistenta. Nedůvěřoval mu.
9	S	Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.	Koperník veřejně podporoval Galileo. Věřil, že Země se otáčí kolem Slunce.	Galileo veřejně podporoval Koperníka. Věřil, že Země se otáčí kolem Slunce.	Galileem byl veřejně podporován Koperník. Věřil, že Země se otáčí kolem Slunce.
10	V	Michael Schumacher porazil Fernanda Alonsa. Jeho fanoušci byli překvapeni.	Michael Schumacher porazil Fernanda Alonsa. Jeho fanoušci byli překvapeni.	Michael Schumacher porazil Fernanda Alonsa. Jeho fanoušci byli překvapeni.	Michael Schumacher porazil Fernanda Alonsa. Jeho fanoušci byli překvapeni.
11	V	Ředitele rozhlasu pozval ředitel televize do vysílání. Moderoval populární pořad.	Ředitele rozhlasu pozval ředitel televize do vysílání. Moderoval populární pořad.	Ředitele rozhlasu pozval ředitel televize do vysílání. Moderoval populární pořad.	Ředitele rozhlasu pozval ředitel televize do vysílání. Moderoval populární pořad.

12	S	Pan listonoš dnes doručil objednaný časopis. Zpozdil se.	Objednaný časopis dnes doručil pan listonoš. Zpozdil se.	Pan listonoš dnes doručil objednaný časopis. Zpozdil se.	Panem listonošem byl dnes doručen objednaný časopis. Zpozdil se.
13	S	Liberální voliči podporují demokratické kandidáty. Bojují za občanské svobody, volný trh a zmírnění chudoby.	Demokratické kandidáty podporují liberální voliči. Bojují za občanské svobody, volný trh a zmírnění chudoby.	Liberální voliči podporují demokratické kandidáty. Bojují za občanské svobody, volný trh a zmírnění chudoby.	Liberálními voliči jsou podporováni demokratičtí kandidáti. Bojují za občanské svobody, volný trh a zmírnění chudoby.
14	V	Autosalon propaguje nový vůz. Jeho úspěšnost závisí na reklamě.	Autosalon propaguje nový vůz. Jeho úspěšnost závisí na reklamě.	Autosalon propaguje nový vůz. Jeho úspěšnost závisí na reklamě.	Autosalon propaguje nový vůz. Jeho úspěšnost závisí na reklamě.
15	S	Palice odrazily útok seker. Jejich zástupci bojovali statečně.	Útok seker odrazily palice. Jejich zástupci bojovali statečně.	Palice odrazily útok seker. Jejich zástupci bojovali statečně.	Palicemi byl odražen útok seker. Jejich zástupci bojovali statečně.
16	S	Prezidenta Václava Klause po deseti letech střídá Miloš Zeman. Jeho příznivci se na inauguraci sjíždějí z celé země.	Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země.	Milošem Zemanem je po deseti letech střídán prezident Václav Klaus. Jeho příznivci se na inauguraci sjíždějí z celé země.	Miloš Zeman po deseti letech střídá prezidenta Václava Klause. Jeho příznivci se na inauguraci sjíždějí z celé země.
17	V	Vánočka šíří po celém bytě vůni. Není snadné jí odolat.	Vánočka šíří po celém bytě vůni. Není snadné jí odolat.	Vánočka šíří po celém bytě vůni. Není snadné jí odolat.	Vánočka šíří po celém bytě vůni. Není snadné jí odolat.
18	V	Postřelenou dívku převezli do jiné nemocnice. Je lépe hlídána.	Postřelenou dívku převezli do jiné nemocnice. Je lépe hlídána.	Postřelenou dívku převezli do jiné nemocnice. Je lépe hlídána.	Postřelenou dívku převezli do jiné nemocnice. Je lépe hlídána.
19	V	Řada odborníků lobbuje za levný vůz. To je to, co Škodě chybí.	Řada odborníků lobbuje za levný vůz. To je to, co Škodě chybí.	Řada odborníků lobbuje za levný vůz. To je to, co Škodě chybí.	Řada odborníků lobbuje za levný vůz. To je to, co Škodě chybí.
20	S	Snaživou kolegyni netrpělivě očekává paní doktorka. Její pacient není připraven na operaci.	Paní doktorka netrpělivě očekává snaživou kolegyni. Její pacient není připraven na operaci.	Paní doktorkou je netrpělivě očekávána snaživou kolegyně. Její pacient není připraven na operaci.	Paní doktorka netrpělivě očekává snaživou kolegyni. Její pacient není připraven na operaci.
21	V	TOP 09 je nyní v Poslanecké sněmovně. O její budoucnost se neobávám.	TOP 09 je nyní v Poslanecké sněmovně. O její budoucnost se neobávám.	TOP 09 je nyní v Poslanecké sněmovně. O její budoucnost se neobávám.	TOP 09 je nyní v Poslanecké sněmovně. O její budoucnost se neobávám.

22	S	Aničku pochválila paní učitelka. Měla radost, že se jí cvičení povedlo.	Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.	Paní učitelkou byla pochválena Anička. Měla radost, že se jí cvičení povedlo.	Paní učitelka pochválila Aničku. Měla radost, že se jí cvičení povedlo.
23	V	Populární zpěvačka byla oceněna cenou Grammy. Fanoušci jsou z ní nadšení.	Populární zpěvačka byla oceněna cenou Grammy. Fanoušci jsou z ní nadšení.	Populární zpěvačka byla oceněna cenou Grammy. Fanoušci jsou z ní nadšení.	Populární zpěvačka byla oceněna cenou Grammy. Fanoušci jsou z ní nadšení.
24	S	Karla srdečně přivítal pan domácí. Půjčil mu peníze.	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.	Panem domácím byl srdečně přivítán Karel. Půjčil mu peníze.	Pan domácí srdečně přivítal Karla. Půjčil mu peníze.
25	V	Evropská vlajka zavlála na Pražském hradě. Zdobí pražské panorama.	Evropská vlajka zavlála na Pražském hradě. Zdobí pražské panorama.	Evropská vlajka zavlála na Pražském hradě. Zdobí pražské panorama.	Evropská vlajka zavlála na Pražském hradě. Zdobí pražské panorama.
26	S	Dvacet nových zaměstnanců přijali minulý měsíc do naší firmy personalisté. Jejich pracovitost je nepřekonatelná.	Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.	Personalisty bylo minulý měsíc do naší firmy přijato dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.	Personalisté minulý měsíc přijali do naší firmy dvacet nových zaměstnanců. Jejich pracovitost je nepřekonatelná.
27	V	Všechny důkazy prozkoumávají kriminalisté. Poskytují srovnání s databází.	Všechny důkazy prozkoumávají kriminalisté. Poskytují srovnání s databází.	Všechny důkazy prozkoumávají kriminalisté. Poskytují srovnání s databází.	Všechny důkazy prozkoumávají kriminalisté. Poskytují srovnání s databází.
28	V	Tahle knížka nemá co dělat v rubrice <i>Tohle čtu</i> . Nic ke čtení v ní není.	Tahle knížka nemá co dělat v rubrice <i>Tohle čtu</i> . Nic ke čtení v ní není.	Tahle knížka nemá co dělat v rubrice <i>Tohle čtu</i> . Nic ke čtení v ní není.	Tahle knížka nemá co dělat v rubrice <i>Tohle čtu</i> . Nic ke čtení v ní není.
29	S	Petr řídí Volkswagen. Jezdí rychle.	Petrovi se dobře řídí Volkswagen. Jezdí rychle.	Petr řídí Volkswagen. Jezdí rychle.	Volkswagen řídí Petr. Jezdí rychle.
30	V	Češi chtějí vidět Jaromíra Jágra. Jejich touha vyprodala O2 arénu.	Češi chtějí vidět Jaromíra Jágra. Jejich touha vyprodala O2 arénu.	Češi chtějí vidět Jaromíra Jágra. Jejich touha vyprodala O2 arénu.	Češi chtějí vidět Jaromíra Jágra. Jejich touha vyprodala O2 arénu.
31	S	Pittsburští Tučňáci odrazili newyorské Rangery. Jejich fanoušci byli překvapeni.	Pittsburskými Tučňáky byli odraženi newyorští Rangeři. Jejich fanoušci byli překvapeni.	Pittsburští Tučňáci odrazili newyorské Rangery. Jejich fanoušci byli překvapeni.	Newyorské Rangery odrazili pittsburští Tučňáci. Jejich fanoušci byli překvapeni.
32	V	Babička bezvýhradně miluje svou vnučku. Je hodná.	Babička bezvýhradně miluje svou vnučku. Je hodná.	Babička bezvýhradně miluje svou vnučku. Je hodná.	Babička bezvýhradně miluje svou vnučku. Je hodná.

33	S	Policie vyšetřuje korupci. Její vliv ve veřejném prostoru je nezpochybnitelný.	Policií je vyšetřována korupce. Její vliv ve veřejném prostoru je nezpochybnitelný.	Policie vyšetřuje korupci. Její vliv ve veřejném prostoru je nezpochybnitelný.	Korupci vyšetřuje policie. Její vliv ve veřejném prostoru je nezpochybnitelný.
34	V	Pavel reklamoval nový foťák. Nefotil.	Pavel reklamoval nový foťák. Nefotil.	Pavel reklamoval nový foťák. Nefotil.	Pavel reklamoval nový foťák. Nefotil.
35	VN	Norové odmítli Evropskou unii. Teď jí dávají Nobelovu cenu míru.	Norové odmítli Evropskou unii. Teď jí dávají Nobelovu cenu míru.	Norové odmítli Evropskou unii. Teď jí dávají Nobelovu cenu míru.	Norové odmítli Evropskou unii. Teď jí dávají Nobelovu cenu míru.
36	V	Podél řeky se má táhnout promenáda. Ta zhorší dopad povodní.	Podél řeky se má táhnout promenáda. Ta zhorší dopad povodní.	Podél řeky se má táhnout promenáda. Ta zhorší dopad povodní.	Podél řeky se má táhnout promenáda. Ta zhorší dopad povodní.
37	S	Pořadatel nejprve uvítal na konferenci zahraničního profesora. Měl radost, že se potkali.	Pořadatelem byl nejprve na konferenci uvítán zahraniční profesor. Měl radost, že se potkali.	Pořadatel nejprve uvítal na konferenci zahraničního profesora. Měl radost, že se potkali.	Zahraničního profesora nejprve uvítal na konferenci pořadatel. Měl radost, že se potkali.
38	VN	Do České filharmonie se vrací bývalý dirigent. Jeho dva koncerty vyvolaly kladné ohlasy.	Do České filharmonie se vrací bývalý dirigent. Jeho dva koncerty vyvolaly kladné ohlasy.	Do České filharmonie se vrací bývalý dirigent. Jeho dva koncerty vyvolaly kladné ohlasy.	Do České filharmonie se vrací bývalý dirigent. Jeho dva koncerty vyvolaly kladné ohlasy.
39	V	Ve svém novém románu čerpá Michal Viewegh z bulváru. Velice ho nesnáší.	Ve svém novém románu čerpá Michal Viewegh z bulváru. Velice ho nesnáší.	Ve svém novém románu čerpá Michal Viewegh z bulváru. Velice ho nesnáší.	Ve svém novém románu čerpá Michal Viewegh z bulváru. Velice ho nesnáší.
40	S	Národní knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	Národní knihovnou byla obdarována fakulta. Bylo to pro ni výhodné.	Národní knihovna darovala knihy fakultě. Bylo to pro ni výhodné.	Fakultě darovala knihy Národní knihovna. Bylo to pro ni výhodné.
41	V	Na svých účtech uložte všechny úspory. Jejich bezpečnost zaručuje banka.	Na svých účtech uložte všechny úspory. Jejich bezpečnost zaručuje banka.	Na svých účtech uložte všechny úspory. Jejich bezpečnost zaručuje banka.	Na svých účtech uložte všechny úspory. Jejich bezpečnost zaručuje banka.
42	V	Ekonomická analýza potvrzuje správnost nově vytvořené strategie. Sklízí pochvaly.	Ekonomická analýza potvrzuje správnost nově vytvořené strategie. Sklízí pochvaly.	Ekonomická analýza potvrzuje správnost nově vytvořené strategie. Sklízí pochvaly.	Ekonomická analýza potvrzuje správnost nově vytvořené strategie. Sklízí pochvaly.

43	S	Ronaldem McDonalDEM byl letos jmenován nový generální ředitel Alexandr Schramm. Jeho prozřetelnost zachránila firmě reputaci.	Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.	Nového generálního ředitele Alexandra Schramma jmenoval letos Ronald McDonald. Jeho prozřetelnost zachránila firmě reputaci.	Ronald McDonald letos jmenoval Alexandra Schramma novým generálním ředitelem. Jeho prozřetelnost zachránila firmě reputaci.
44	S	Fotbalisty Bohemians byli na zápas pozváni kluci. Holky je obdivovaly.	Fotbalisté Bohemians pozvali na zápas kluky. Holky je obdivovaly.	Kluky pozvali na zápas fotbalisté Bohemians. Holky je obdivovaly.	Fotbalisté Bohemians pozvali na zápas kluky. Holky je obdivovaly.
45	VN	Budova nového pavilonu v botanické zahradě bude patřit různým. Ty jsou novým logem zahrady.	Budova nového pavilonu v botanické zahradě bude patřit různým. Ty jsou novým logem zahrady.	Budova nového pavilonu v botanické zahradě bude patřit různým. Ty jsou novým logem zahrady.	Budova nového pavilonu v botanické zahradě bude patřit různým. Ty jsou novým logem zahrady.
46	S	Začínajícím autorem je na veletrhu představován jeho první román. Velký zájem o něj nastartovala televizní reklama.	Začínající autor na veletrhu představuje svůj první román. Velký zájem o něj nastartovala televizní reklama.	Svůj první román představuje na veletrhu začínající autor. Velký zájem o něj nastartovala televizní reklama.	Začínající autor na veletrhu představuje svůj první román. Velký zájem o něj nastartovala televizní reklama.
47	VN	České fotbalisty čeká další bitva. Snad dopadne dobře.	České fotbalisty čeká další bitva. Snad dopadne dobře.	České fotbalisty čeká další bitva. Snad dopadne dobře.	České fotbalisty čeká další bitva. Snad dopadne dobře.
48	V	Jablka musela oloupat Petra. Byla otrávená.	Jablka musela oloupat Petra. Byla otrávená.	Jablka musela oloupat Petra. Byla otrávená.	Jablka musela oloupat Petra. Byla otrávená.
49	S	Známým architektem Gaudím byl navržen moderní dům ve Francouzské čtvrti. Proslavil ho.	Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil ho.	Moderní dům ve Francouzské čtvrti navrhl známý architekt Gaudí. Proslavil ho.	Známý architekt Gaudí navrhl moderní dům ve Francouzské čtvrti. Proslavil ho.
50	S	Tomášem byl na zápas pozván Michal. Byl velký fanoušek fotbalu.	Tomáš pozval na zápas Michala. Byl velký fanoušek fotbalu.	Michala pozval na zápas Tomáš. Byl velký fanoušek fotbalu.	Tomáš pozval na zápas Michala. Byl velký fanoušek fotbalu.

## 10.7 Otázky a možnosti odpovědi

	otázka	odpověď 1	odpověď 2
1	Koho/co získal nový majitel?	obraz Emila Filly	rekord
2	Komu tleskalo publikum ve stoje?	Rogeru Federerovi	Rafaelu Nadalovi
3	Co apeluje na třídní nevraživost?	vize jednotné Ameriky	nový přístup
4	Kdo/co závisí na hlasech voličů?	dozorčí rada	volební výsledky
5	Komu chutnala whiskey?	Barleymu	bratrovi
6	Kdo si domluvil schůzku s kým?	Klára s dlouholetou kolegyní	dlouholetá kolegyně s Klárou
7	Kdo nechtěl odpovídat na otázky?	starosta	moderátor
8	Kdo nedůvěřoval komu?	šéf asistentovi	asistent šéfovi
9	Kdo věřil, že se Země otáčí kolem Slunce?	Koperník	Galileo
10	Čí fanoušci byli překvapeni?	Fernanda Alonsa	Michaela Schumachera
11	Kdo moderoval populární pořad?	ředitel televize	ředitel rozhlasu
12	Kdo/co se zpozdil/o?	objednaný časopis	pan listonoš
13	Kdo bojuje za občanské svobody, volný trh a zmírnění chudoby?	liberální voliči	demokratičtí kandidáti
14	Čí úspěšnost závisí na reklamě?	autosalonu	nového vozu
15	Čí zástupci bojovali statečně?	seker	palic
16	Kdo střídá koho?	prezidenta Václava Klause Miloš Zeman	Miloše Zemana prezident Václav Klaus
17	Čemu není snadné odolat?	vánočce	vůni
18	Kdo/co je lépe hlídáný/é?	postřelená dívka	jiná nemocnice
19	Co chybí Škodě?	levný vůz	řada odborníků
20	Čí pacient není připraven na operaci?	snaživé kolegyně	paní doktorky
21	O budoucnost koho/čeho se autor neobává?	TOP 09	Poslanecké sněmovny
22	Kdo měl radost?	paní učitelka	Anička
23	Z koho/čeho jsou fanoušci nadšení?	z populární zpěvačky	z ceny Grammy
24	Kdo půjčil peníze komu?	pan domácí Karlovi	Karel panu domácímu
25	Co zdobí pražské panorama?	evropská vlajka	Pražský hrad
26	Čí pracovitost je nepřekonatelná?	personalistů	dvaceti nových zaměstnanců
27	Kdo/co poskytuje srovnání s databází?	všechny důkazy	kriminalisté
28	V čem není nic ke čtení?	v rubrice <i>Tohle čtu</i>	v téhle knížce
29	Kdo/co jezdí rychle?	Petr	Volkswagen
30	Čí touha vyprodala O2 arénu?	Jaromíra Jágra	Čechů
31	Čí fanoušci byli překvapeni?	pittsburských Tučňáků	newyorských Rangerů
32	Kdo je hodný?	vnučka	babička
33	Čí vliv ve veřejném prostoru je nezpochybnitelný?	korupce	policie
34	Kdo/co nefotil/o?	nový foťák	Pavel
35	Kdo dává Nobelovu cenu míru komu?	Norové Evropské unii	Evropská unie Norům
36	Co zhorší dopad povodní?	promenáda	řeka

37	Kdo měl radost, že se potkali?	pořadatel	zahraniční profesor
38	Čí dva koncerty vyvolaly kladné ohlasy?	bývalého dirigenta	České filharmonie
39	Kdo/co nesnáší koho/co?	bulvár Michala Viewegha	Michal Viewegh bulvár
40	Pro koho to bylo výhodné?	pro fakultu	pro Národní knihovnu
41	Čí bezpečnost zaručuje banka?	úctů	úspor
42	Co sklízí pochvaly?	nově vytvořená strategie	ekonomická analýza
43	Čí prozřetelnost zachránila firmě reputaci?	nového generálního ředitele	Ronalda McDonalda
44	Koho holky obdivovaly?	fotbalisty Bohemians	kluky
45	Co je novým logem zahrady?	budova nového pavilonu	růže
46	O koho/co nastartovala zájem televizní reklama?	o začínajícího autora	o jeho první román
47	Kdo/co snad dopadne dobře?	čeští fotbalisté	další bitva
48	Kdo/co byl/o otrávený/é?	Petra	jablka
49	Kdo/co proslavil/o koho/co?	moderní dům architekta Gaudího	architekt Gaudí moderní dům
50	Kdo byl velký fanoušek fotbalu?	Tomáš	Michal

### 10.8 T–F podle nastavení kontextu (chi-square, p-values)

	T > T vs. T > F	F > F vs. F > T
Q1	0.001	1.384e-11
Q2	2.215e-11	0.001
Q3	1.276e-06	0.019
Q4	3.355e-07	4.162e-05

### 10.9 Po–Pt podle nastavení kontextu (chi-square, p-values)

	Po > Po vs. Po > Pt	Pt > Pt vs. Pt > Po
Q1	4.274e-07	2.2e-16
Q2	8.062e-12	4.874e-09
Q3	0.000	0.048
Q4	7.643e-16	6.976e-11